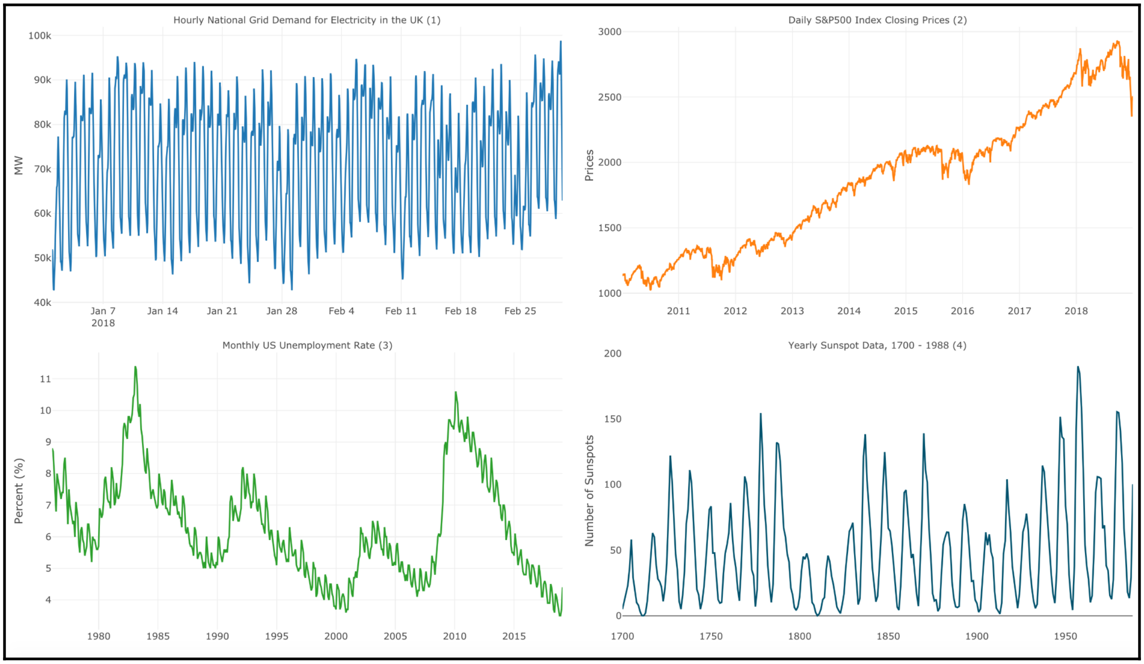
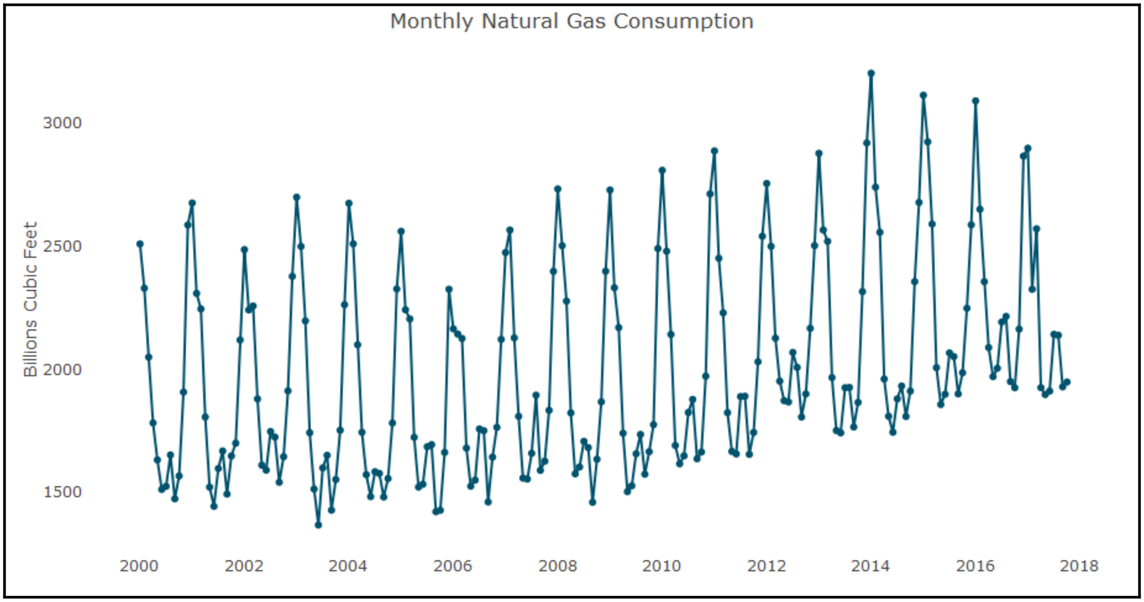
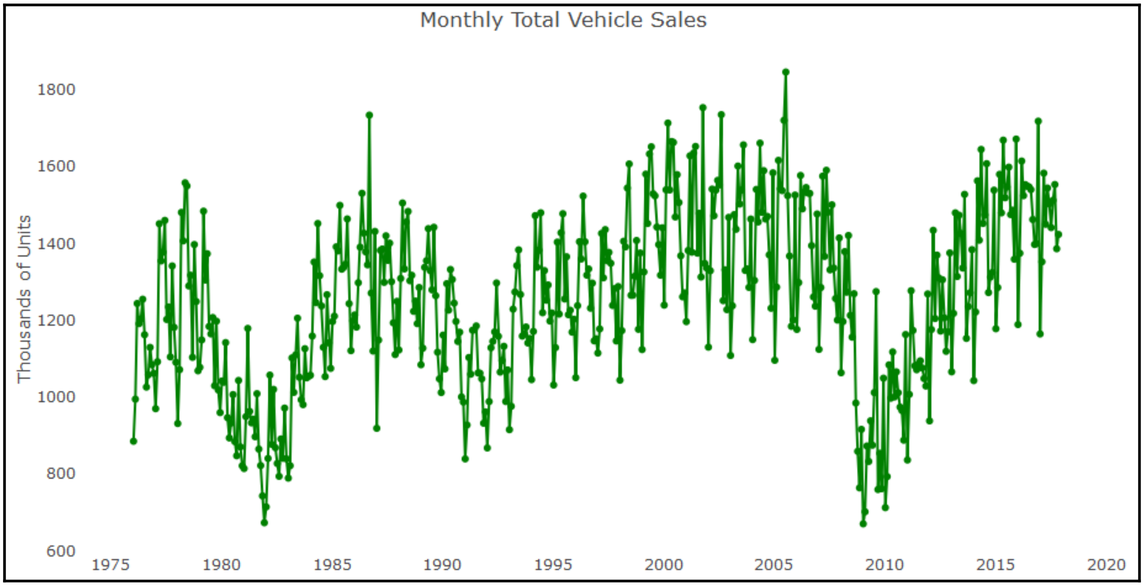
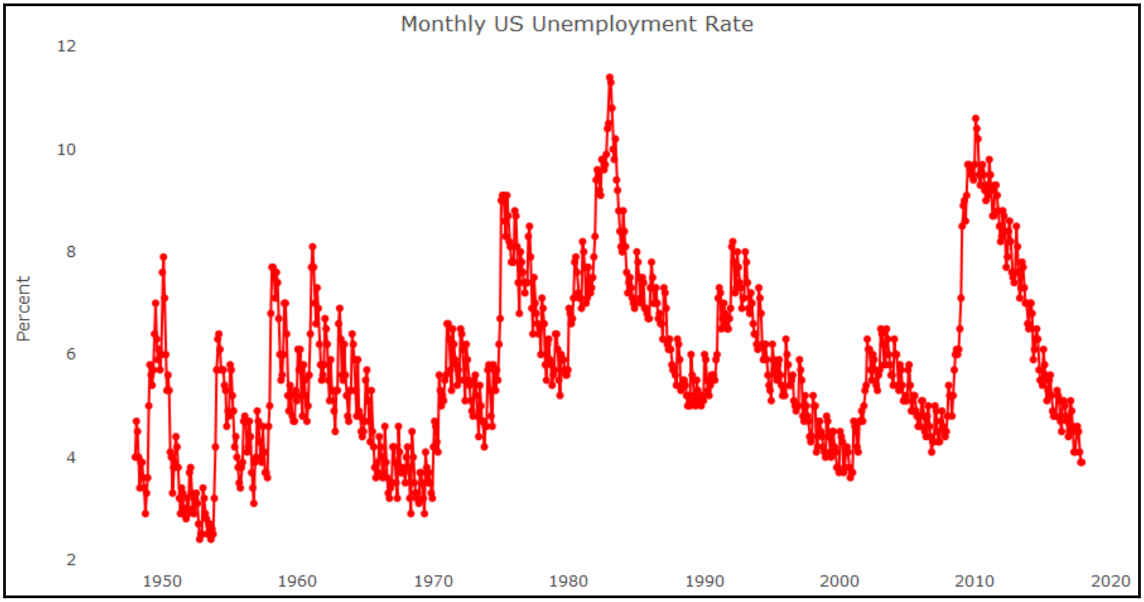
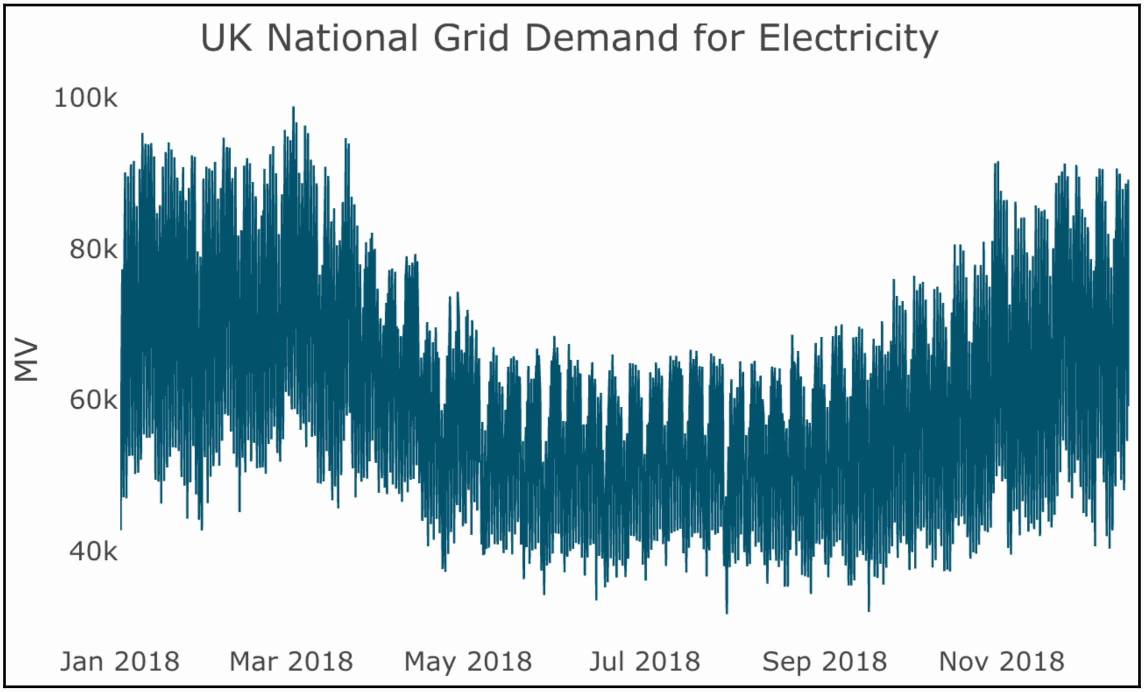
Chapter 1: Introduction to Time Series Analysis and R

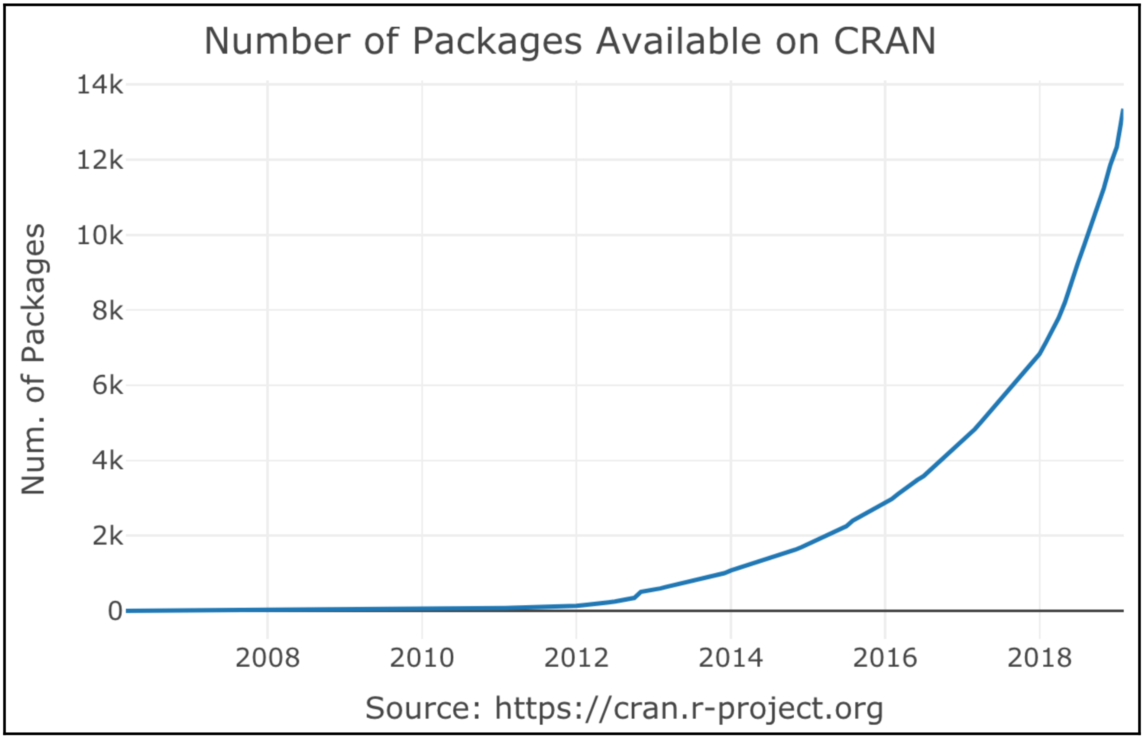



Monthly Natural Gas Consumption
Billions Cubic Feet
3000
2500
2000
1500
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018

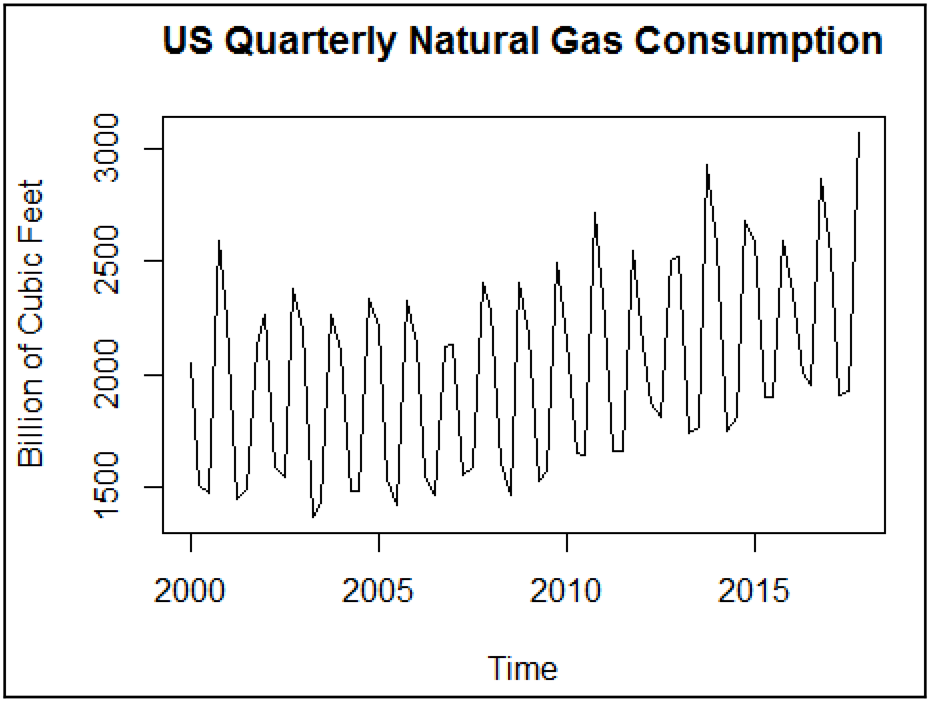

Monthly Total Vehicle Sales
Thousands of Units
1800
1600
1400
1200
1000
800
600
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020


Monthly US Unemployment Rate
Percent
12
10
8
6
4
2
1950
1960
1970
1980
1990
2000
2010
2020

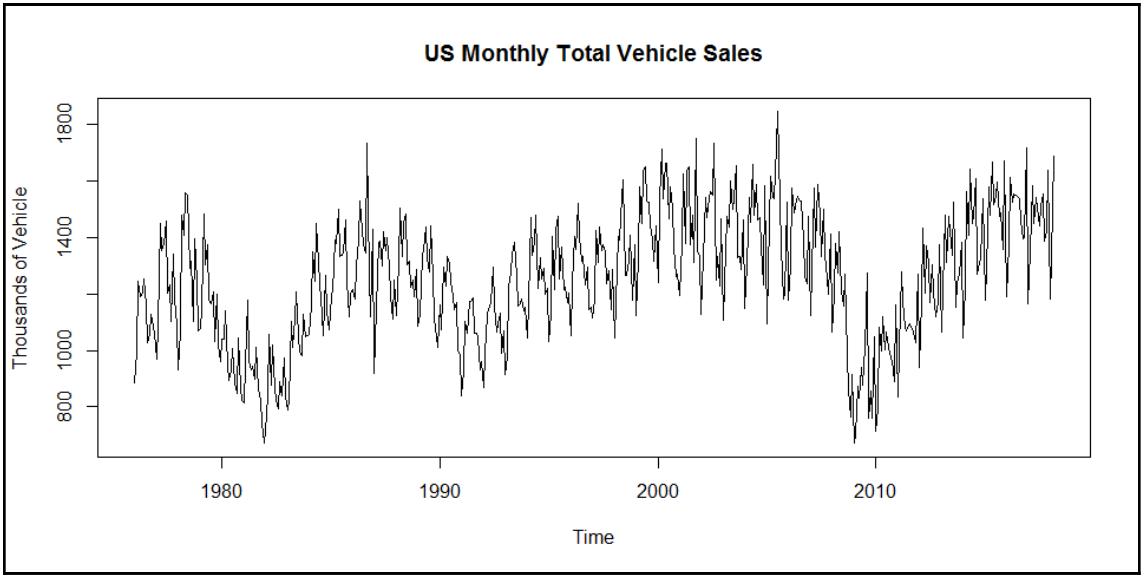

UK National Grid Demand for Electricity
MV
100k
80k
60k
40k
Jan 2018
Mar 2018
May 2018
Jul 2018
Sep 2018
Nov 2018

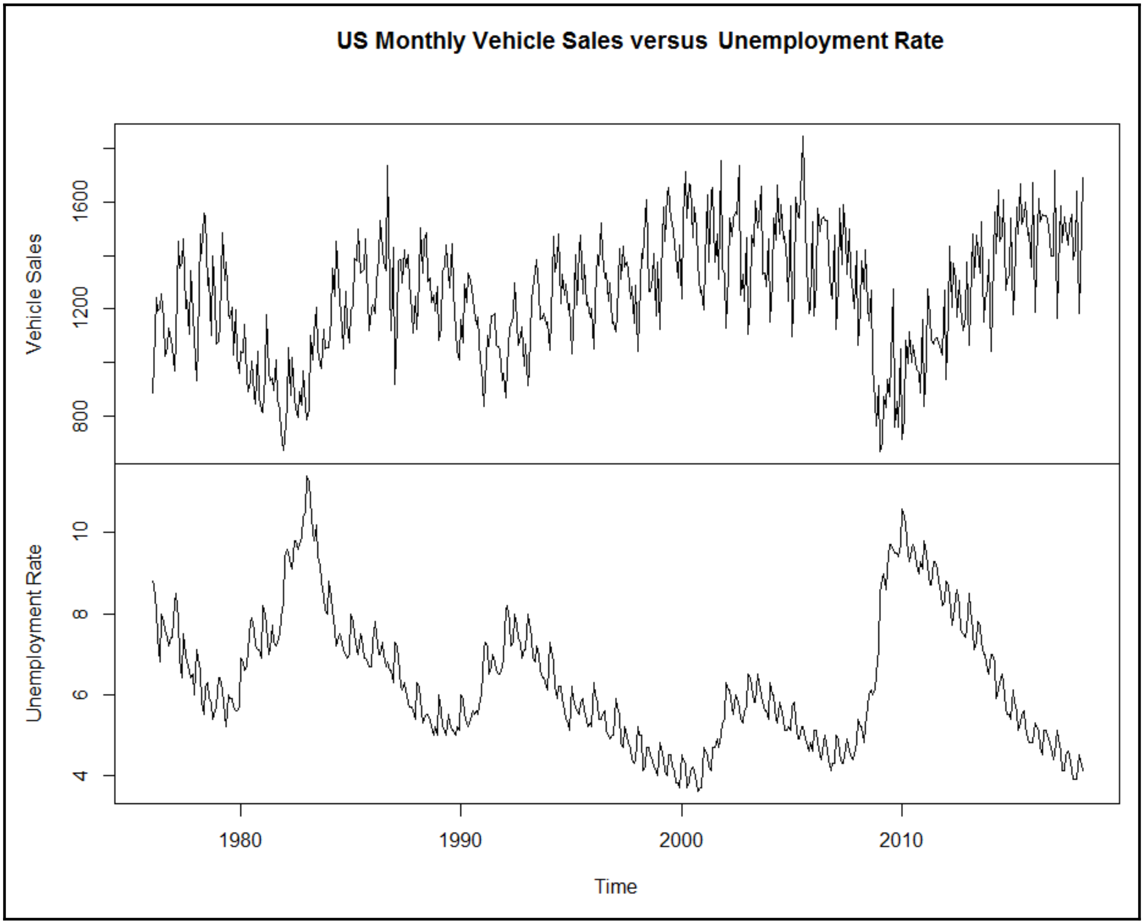

Number of Packages Available on CRAN
14k
12k
10k
8k
6k
4k
2k
0
Num. of Packages
2008
2010
2012
2014
2016
2018
Source: https://cran.r-project.org

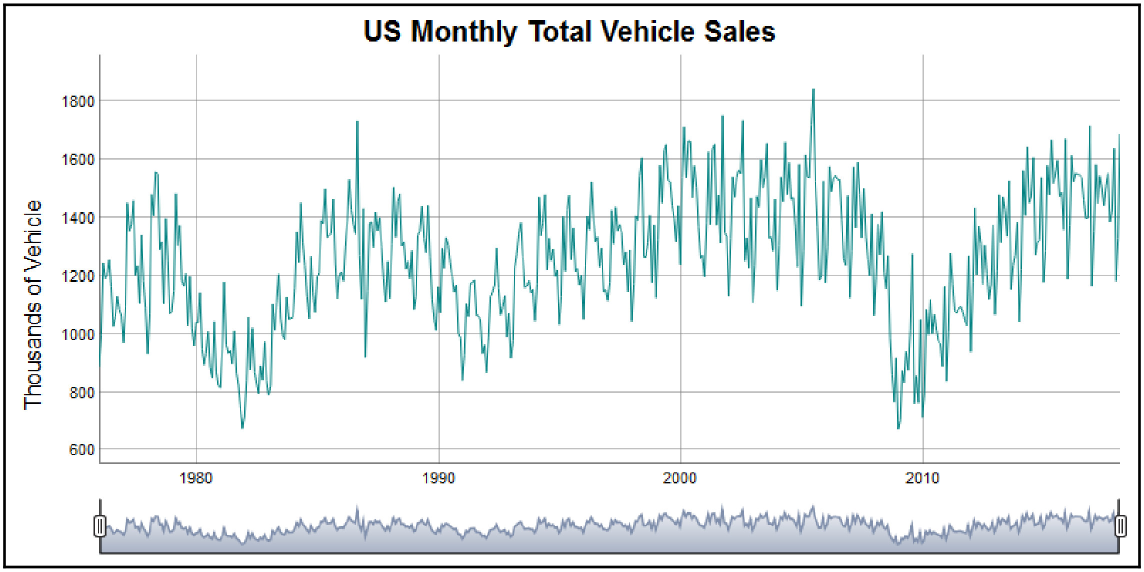
Chapter 3: The Time Series Object

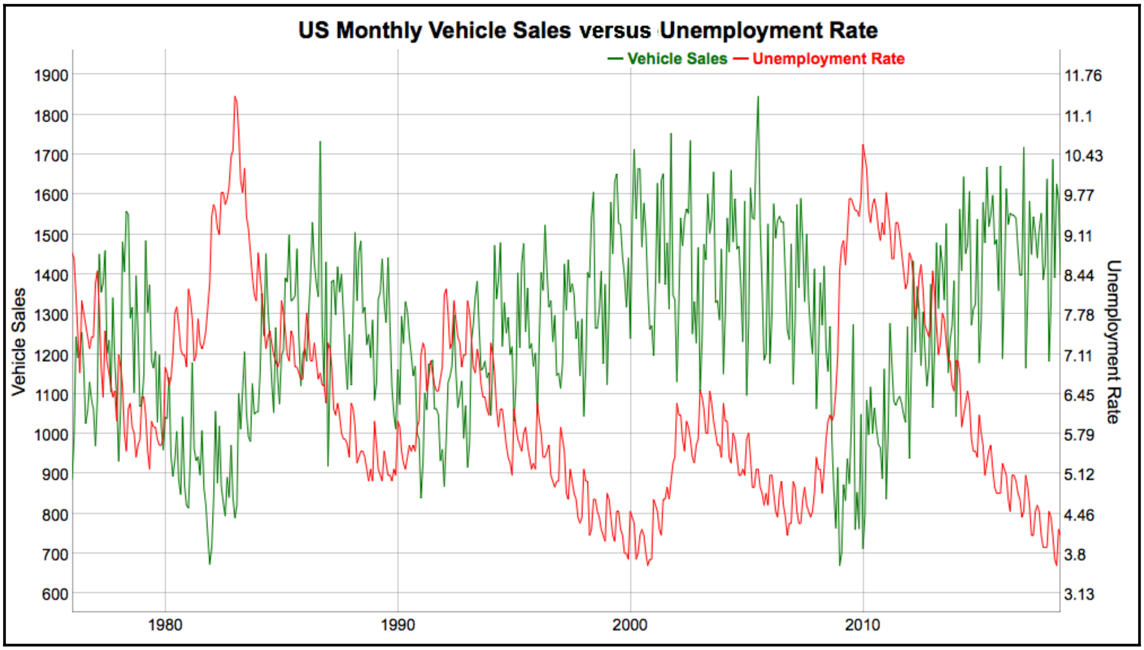


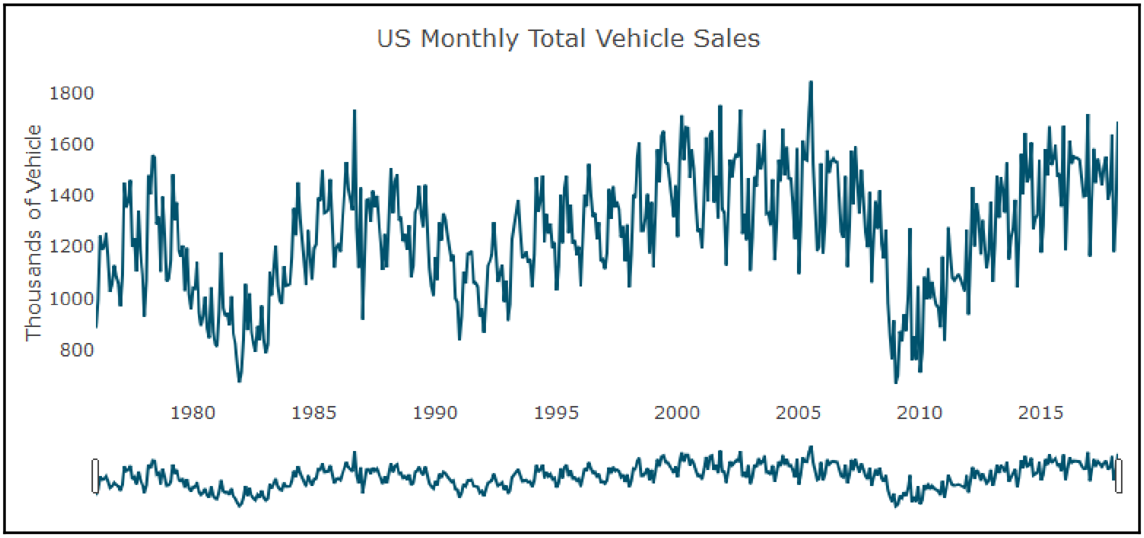
Timestamp	Lag-1	Series	Lead -1
Jan-18	NA	3	5.5
Feb-18	3	5.5	4.7
Mar-18	5.5	4.7	3.6
Apr-18	4.7	3.6	2.5
May-18	3.6	2.5	4.3
Jun-18	2.5	4.3	2.8
Jul-18	4.3	2.8	2.2
Aug-18	2.8	2.2	1.9
Sep-18	2.2	1.9	3.6
Oct-18	1.9	3.6	4.3
Nov-18	3.6	4.3	2.9
Dec-18	4.3	2.9	NA

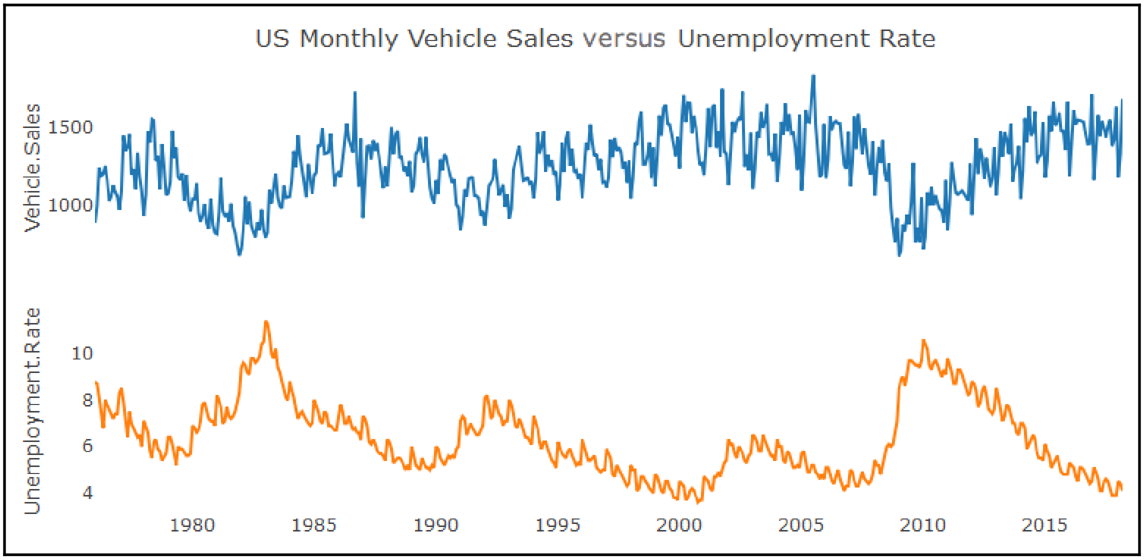

US Monthly Total Vehicle Sales
Thousands of Vehicle
800
1000
1400
1800
1980
1990
2000
2010
Time


US Monthly Vehicle Sales versus Unemployment Rate
Vehicle Sales
800
1200
1600
Unemployment Rate
4
6
8
10
1980
1990
2000
2010
Time

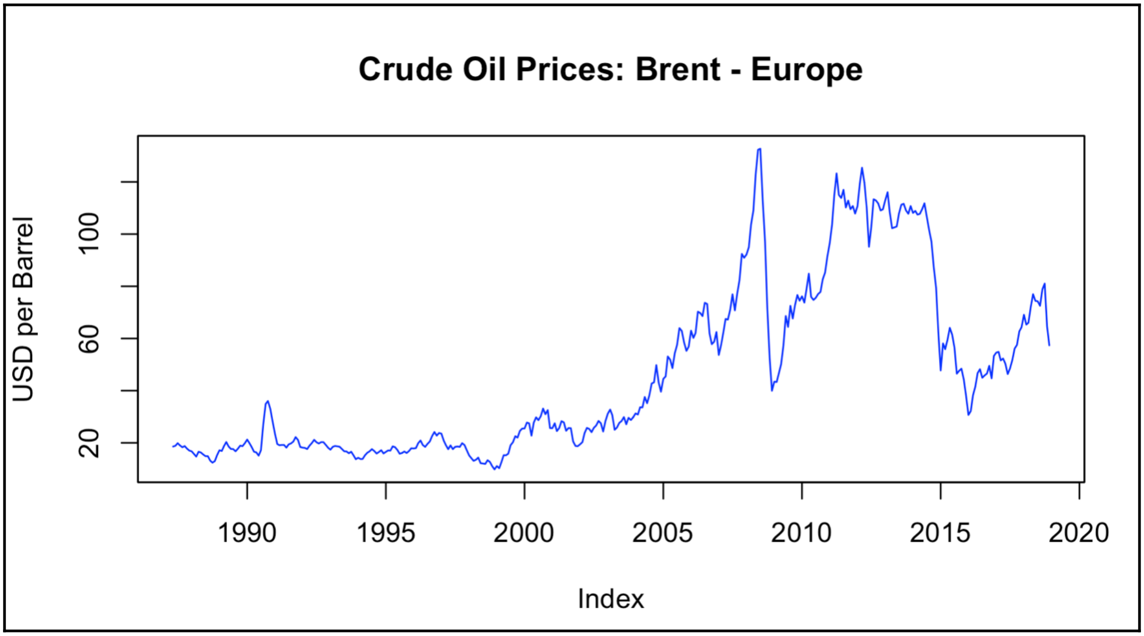

US Monthly Total Vehicle Sales
Thousands of Vehicle
1800
1600
1400
1200
1000
800
600
1980
1990
2000
2010

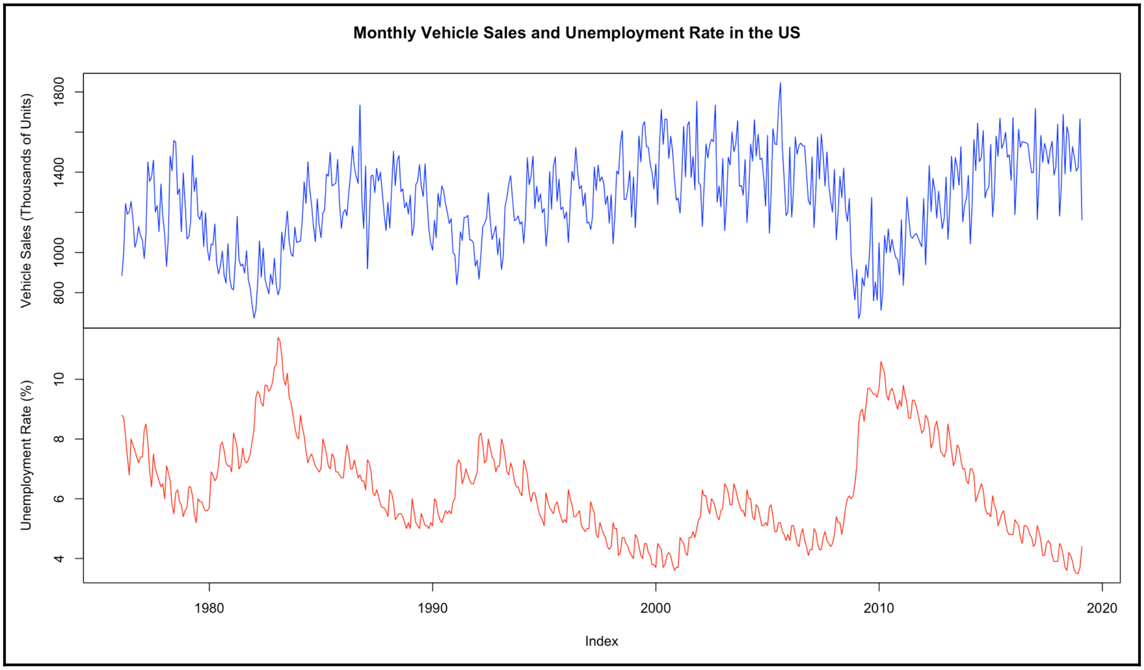

US Monthly Vehicle Sales versus Unemployment Rate
Vehicle Sales
Unemployment Rate
Vehicle Sales
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
11.76
11.1
10.43
9.77
9.11
8.44
7.78
7.11
6.45
5.79
5.12
4.46
3.8
3.13
Unemployment Rate
1980
1990
2000
2010

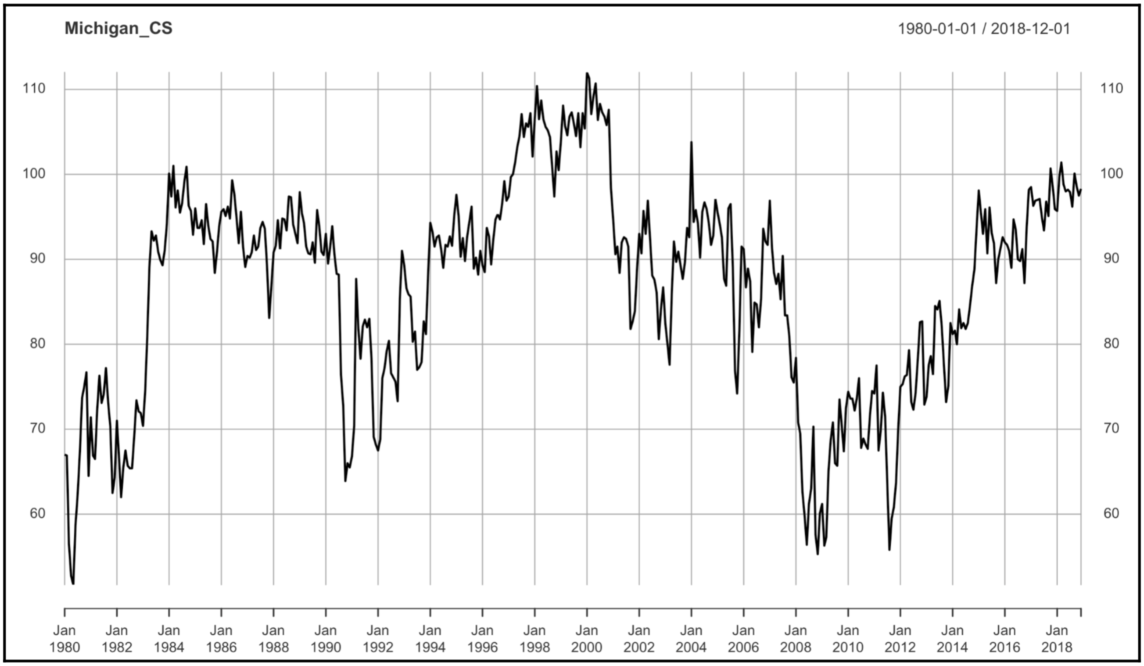

US Monthly Total Vehicle Sales
Thousands of Vehicle
1800
1600
1400
1200
1000
800
1980
1985
1990
1995
2000
2005
2010
2015

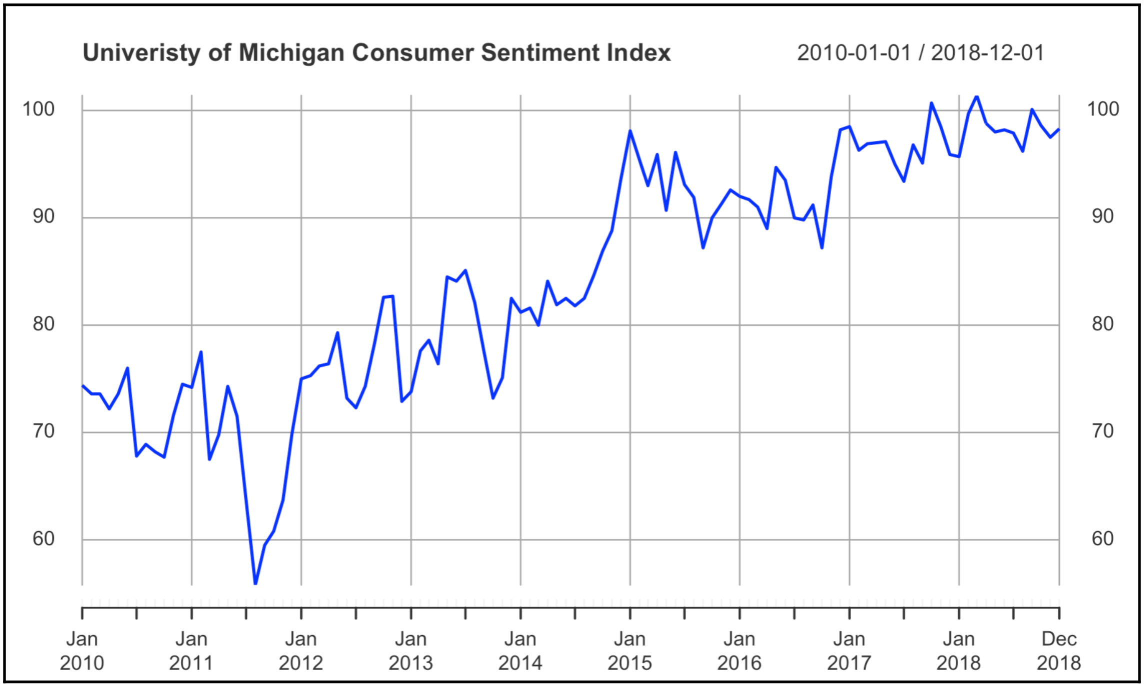

US Monthly Vehicle Sales versus Unemployment Rate
Vehicle.Sales
1500
1000
Unemployment.Rate
10
8
6
4
1980
1985
1990
1995
2000
2005
2010
2015

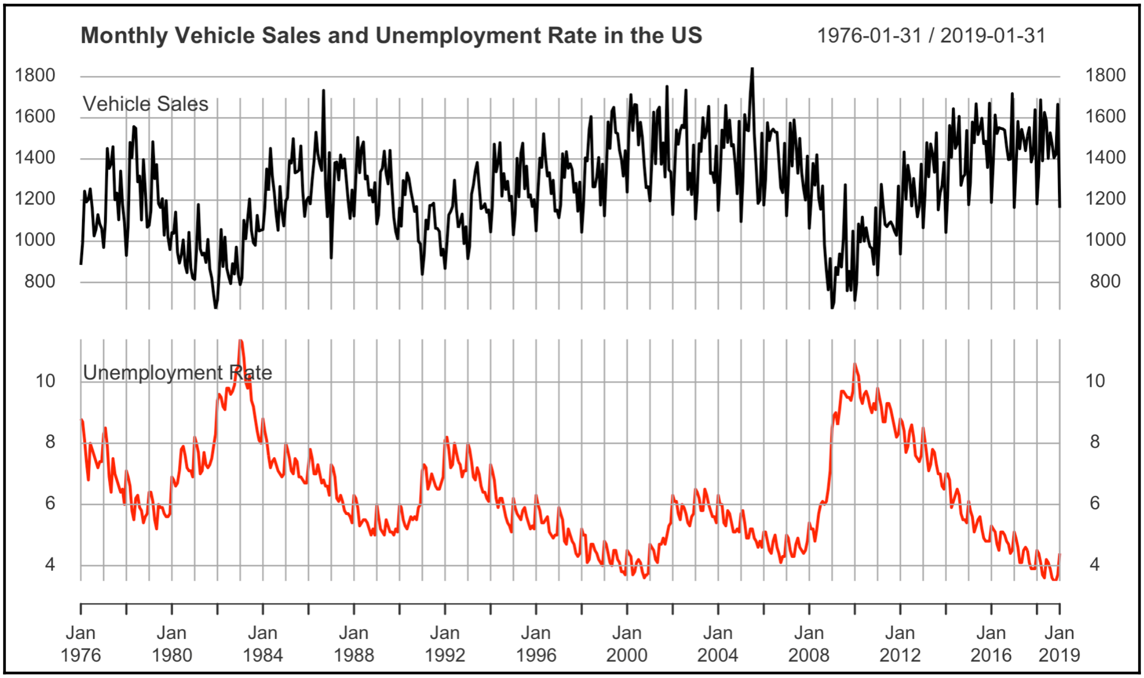
Chapter 4: Working with zoo and xts Objects

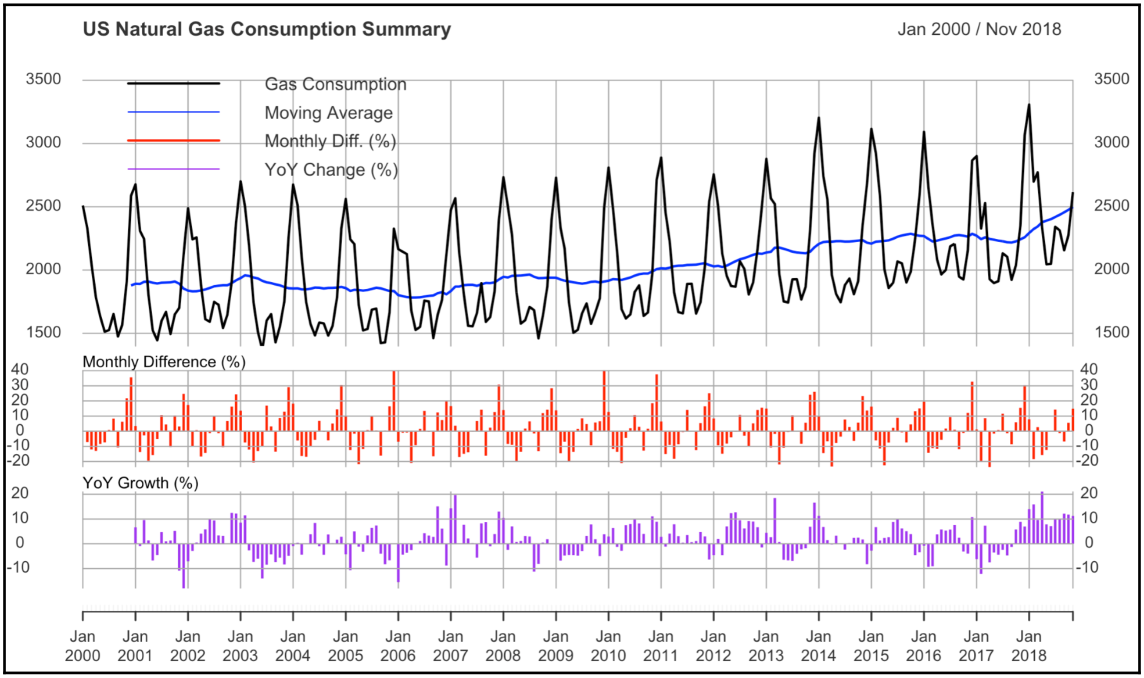



Monthly Vehicle Sales and Unemployment Rate in the US
Vehicle Sales (Thousands of Units)
800
1000
1400
1800
Unemployment Rate (%)
4
6
8
10
1980
1990
2000
2010
2020
Index

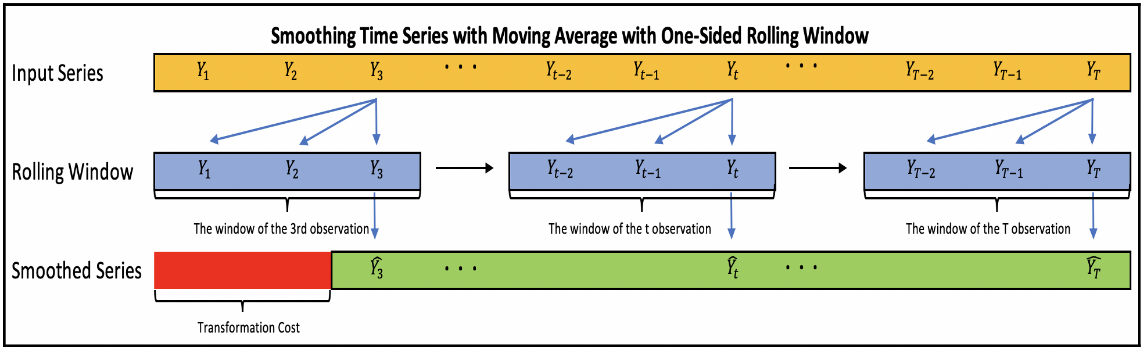

Michigan_CS
1980-01-01 / 2018-12-01
110
100
90
80
70
60
Jan 1980
Jan 1982
Jan 1984
Jan 1986
Jan 1988
Jan 1990
Jan 1992
Jan 1994
Jan 1996
Jan 1998
Jan 2000
Jan 2002
Jan 2004
Jan 2006
Jan 2008
Jan 2010
Jan 2012
Jan 2014
Jan 2016
Jan 2018

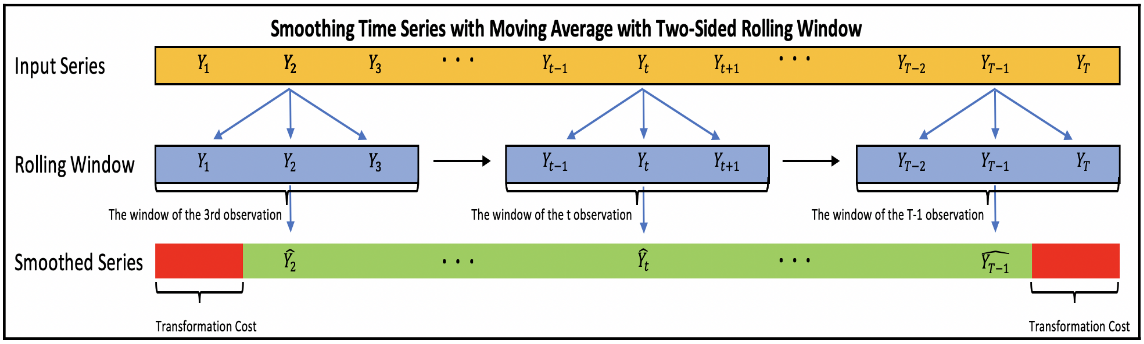

Univeristy of Michigan Consumer Sentiment Index
2010-01-01 / 2018-12-01
100
90
80
70
60
Jan 2010
Jan 2011
Jan 2012
Jan 2013
Jan 2014
Jan 2015
Jan 2016
Jan 2017
Jan 2018
Dec 2018

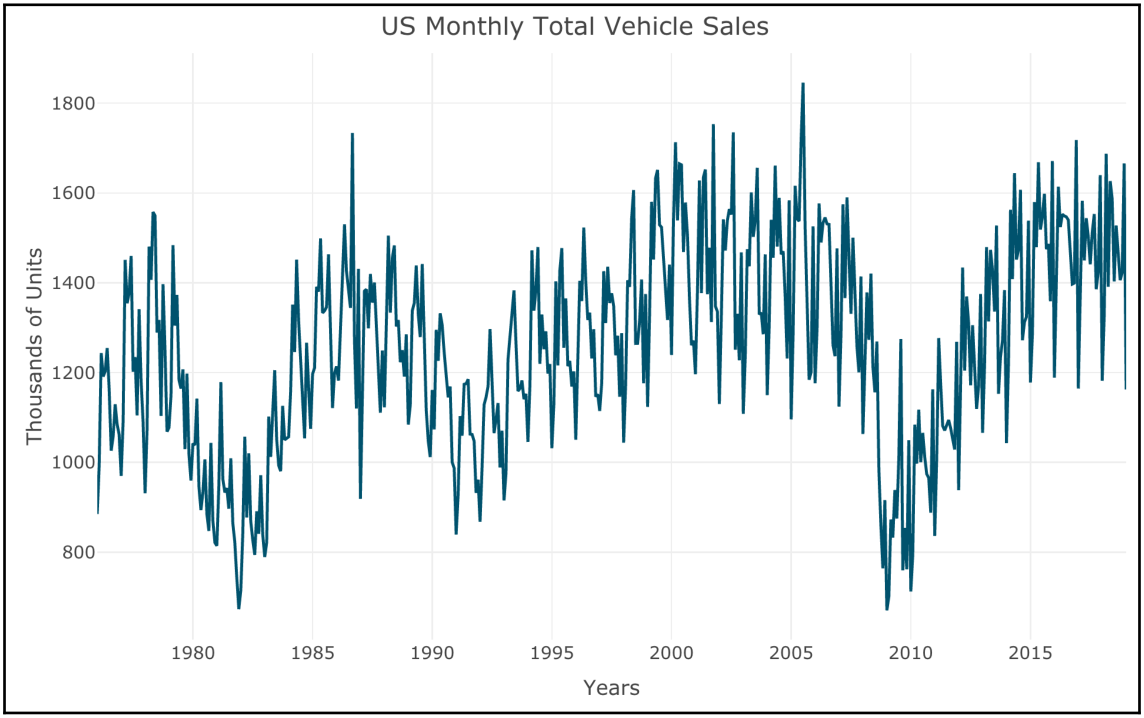

Monthly Vehicle Sales and Unemployment Rate in the US
1976-01-31 / 2019-01-31
Vehicle Sales
1800
1600
1400
1200
1000
800
Unemployment Rate
10
8
6
4
Jan 1976
Jan 1980
Jan 1984
Jan 1988
Jan 1992
Jan 1996
Jan 2000
Jan 2004
Jan 2008
Jan 2012
Jan 2016
Jan 2019

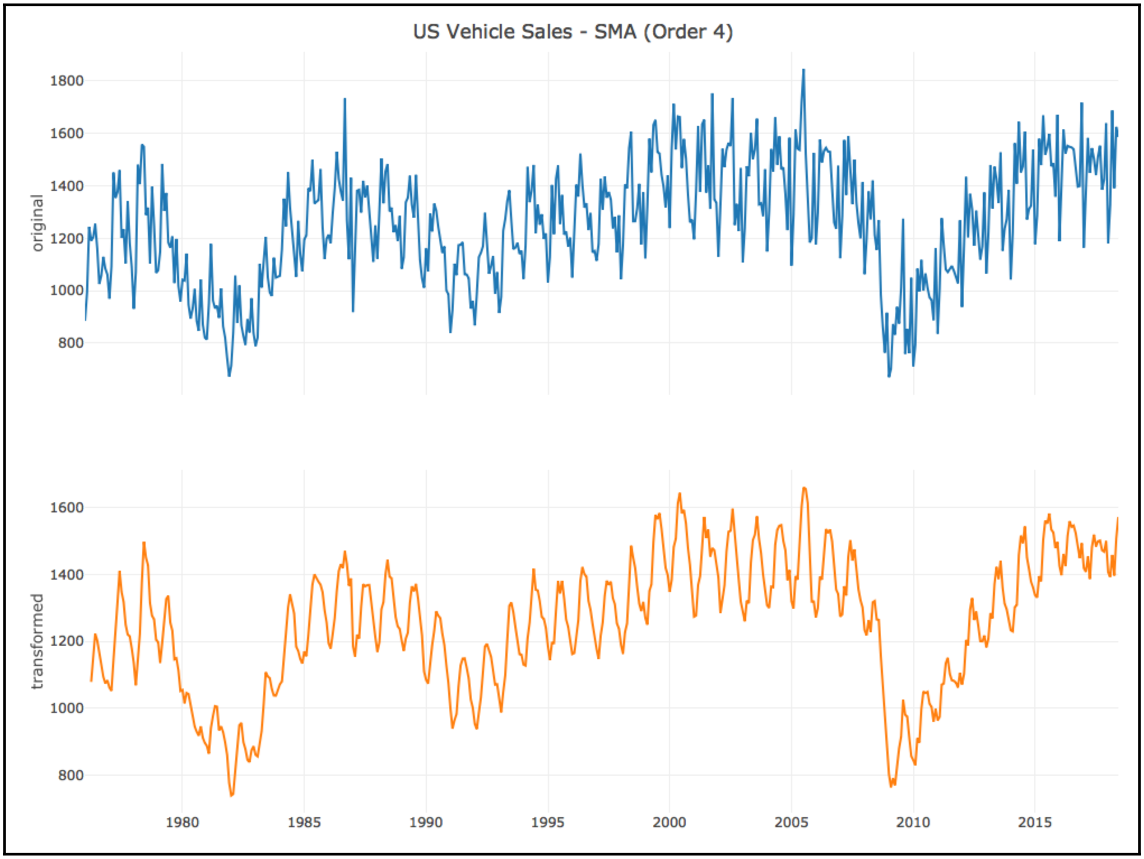

US Natural Gas Consumption Summary
Jan 2000 / Nov 2018
Gas Consumption
Moving Average
Monthly Diff. (%)
YoY Change (%)
3500
3000
2500
2000
1500
Monthly Difference (%)
40
30
20
10
0
-10
-20
YoY Growth (%)
20
10
0
-10
Jan 2000
Jan 2001
Jan 2002
Jan 2003
Jan 2004
Jan 2005
Jan 2006
Jan 2007
Jan 2008
Jan 2009
Jan 2010
Jan 2011
Jan 2012
Jan 2013
Jan 2014
Jan 2015
Jan 2016
Jan 2017
Jan 2018

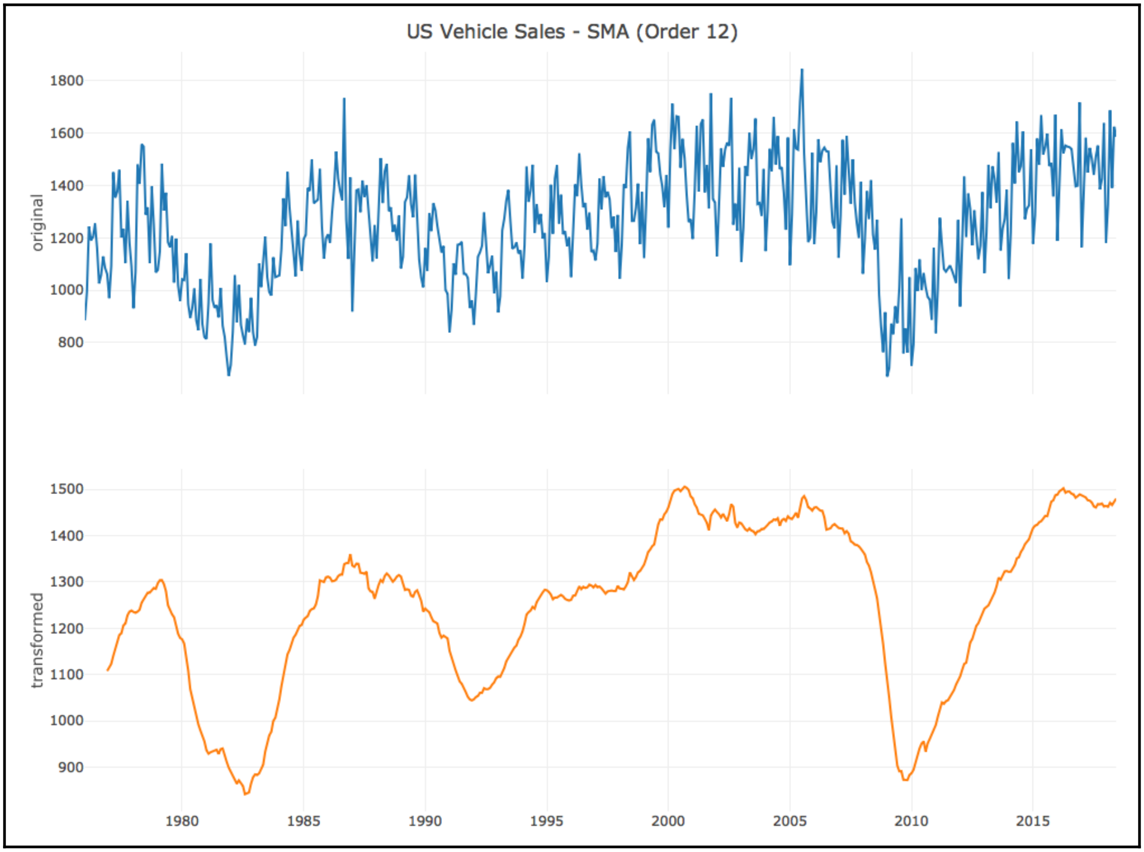
Chapter 5: Decomposition of Time Series Data

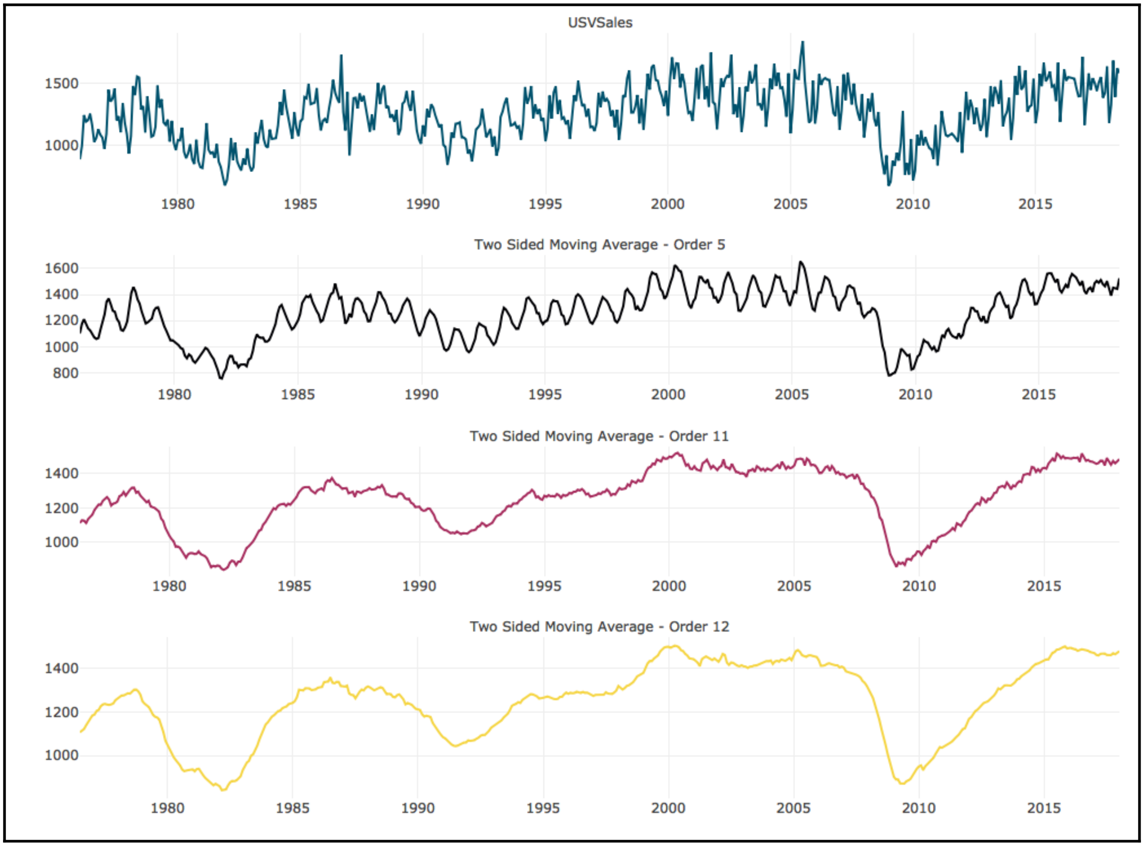


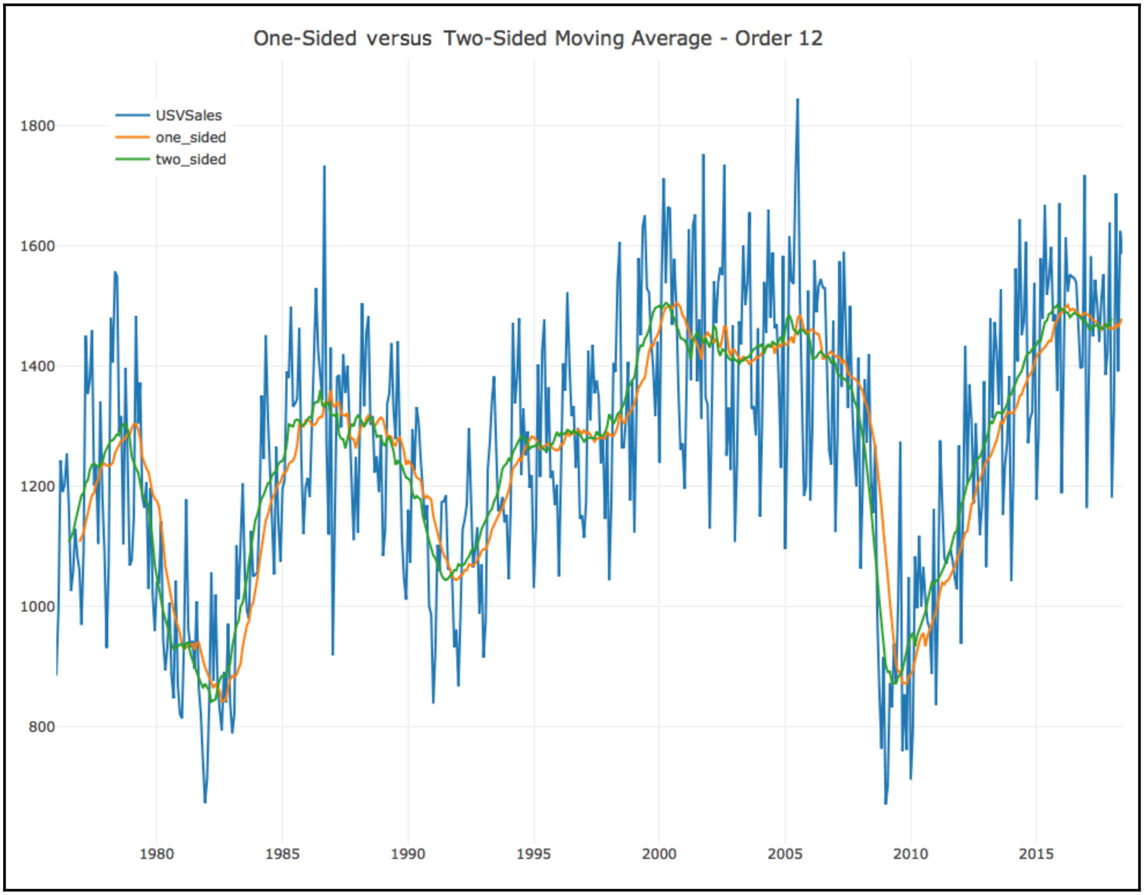


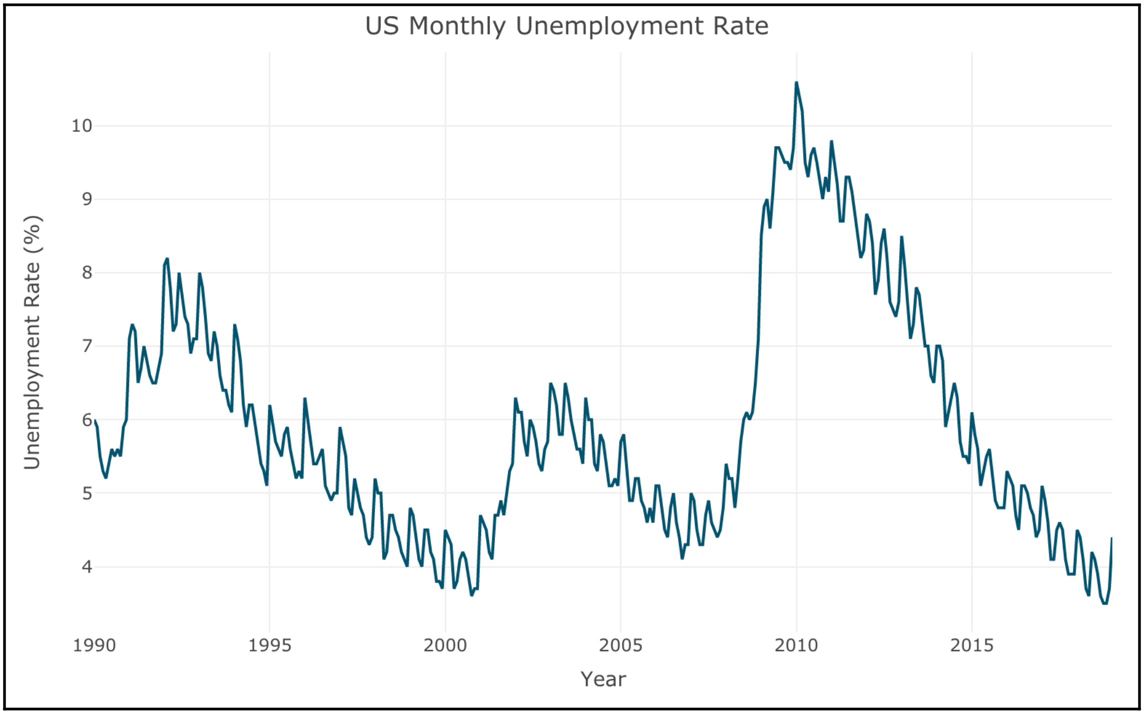

US Monthly Total Vehicle Sales
Thousands of Units
1800
1600
1400
1200
1000
800
1980
1985
1990
1995
2000
2005
2010
2015
Years


US Vehicle Sales - SMA (Order 4)
original
transformed
1800
1600
1400
1200
1000
800
1980
1985
1990
1995
2000
2005
2010
2015

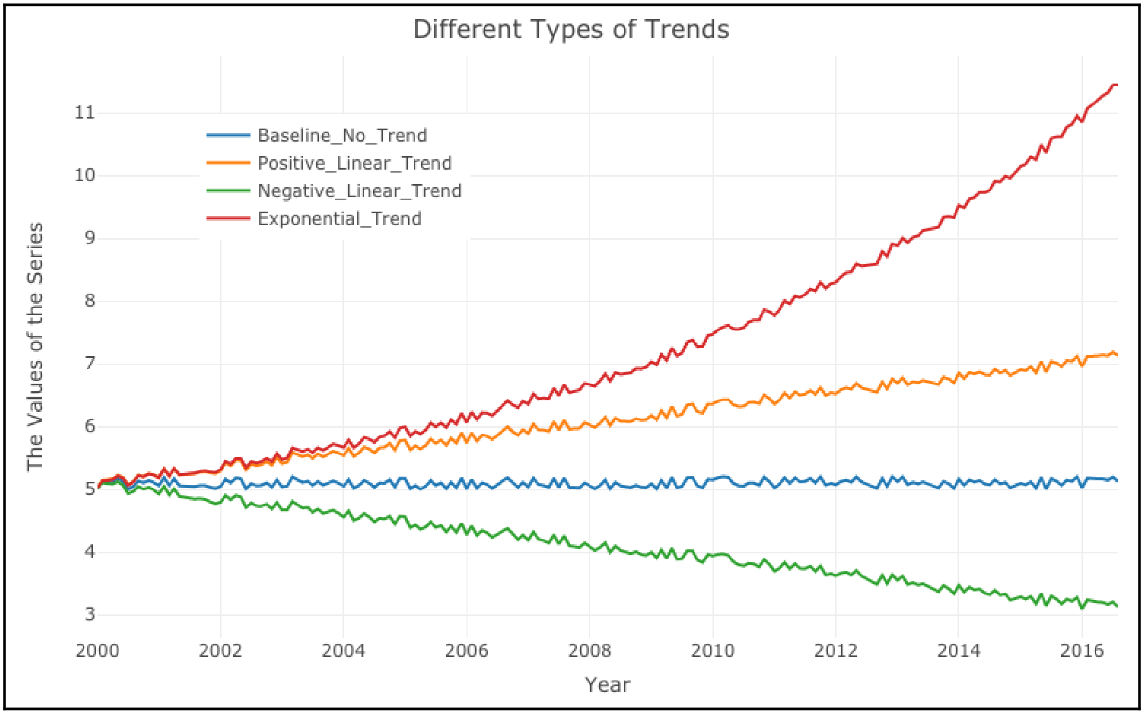

US Vehicle Sales - SMA (Order 12)
original
1800
1600
1400
1200
1000
800
transformed
1500
1400
1300
1200
1100
1000
900
1980
1985
1990
1995
2000
2005
2010
2015

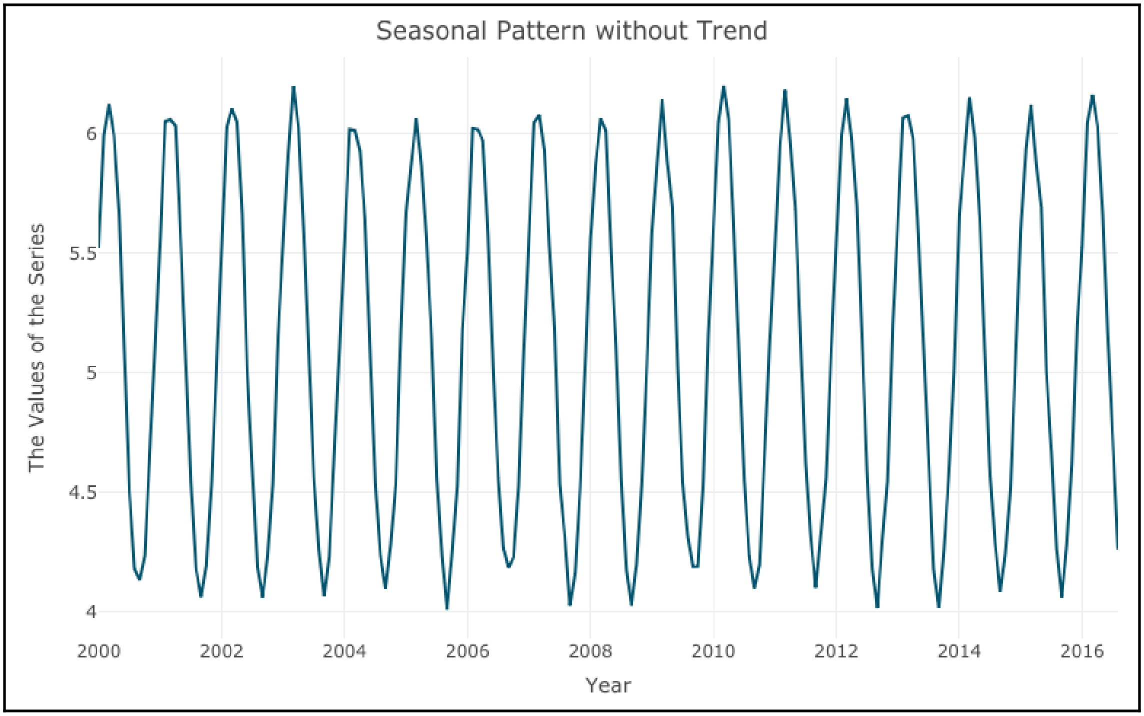

USVSales
1500
1000
1980
1985
1990
1995
2000
2005
2010
2015
Two Sided Moving Average - Order 5
1600
1400
1200
1000
800
Two Sided Moving Average - Order 11
Two Sided Moving Average - Order 12

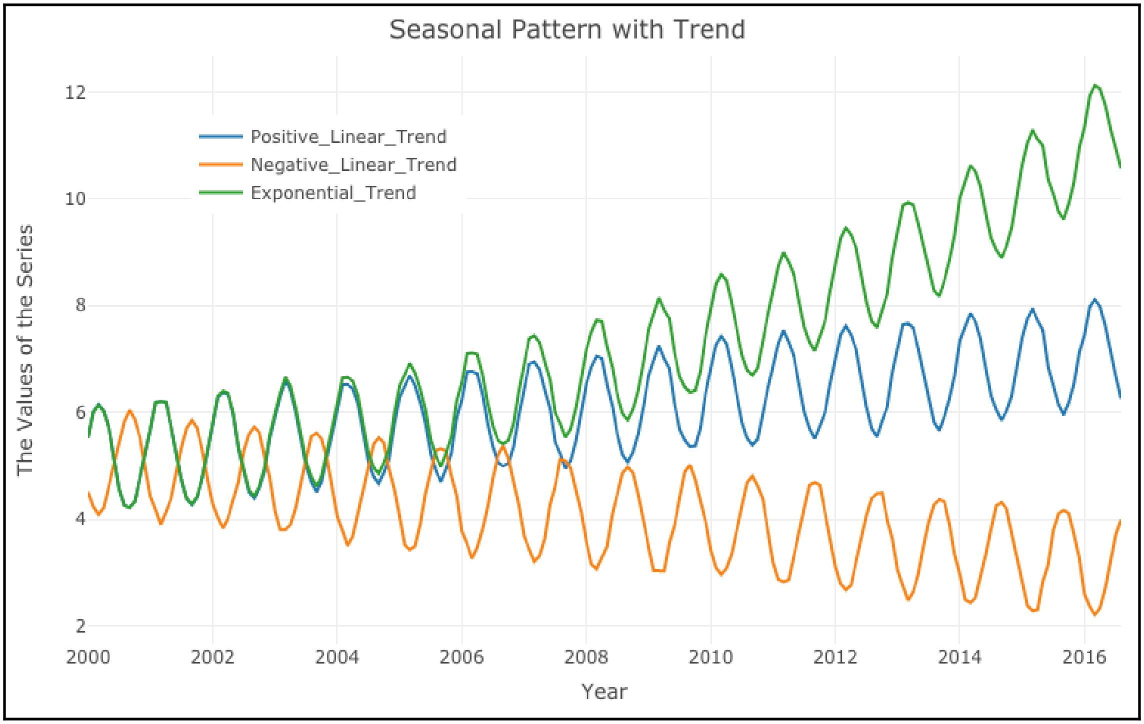

One-Sided versus Two-Sided Moving Average - Order 12
USVSales
one_sided
two_sided
1800
1600
1400
1200
1000
800
1980
1985
1990
1995
2000
2005
2010
2015

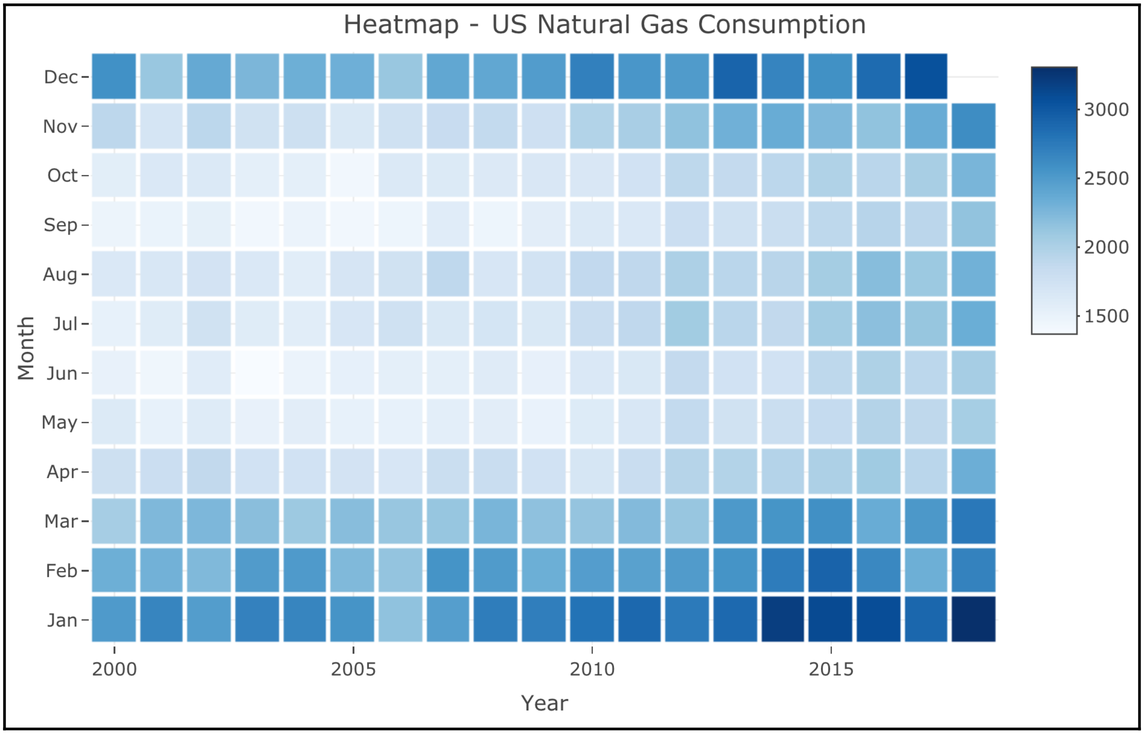


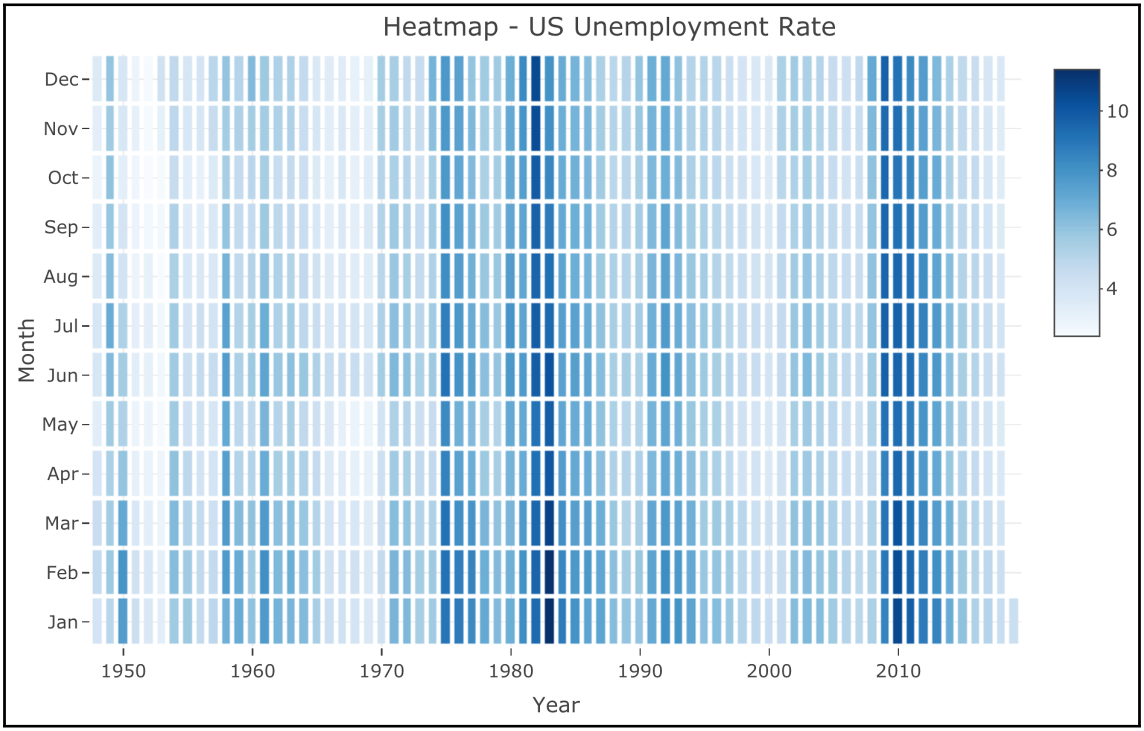
US Monthly Unemployment Rate
Unemployment Rate (%)
10
9
8
7
6
5
4
1990
1995
2000
2005
2010
2015
Year

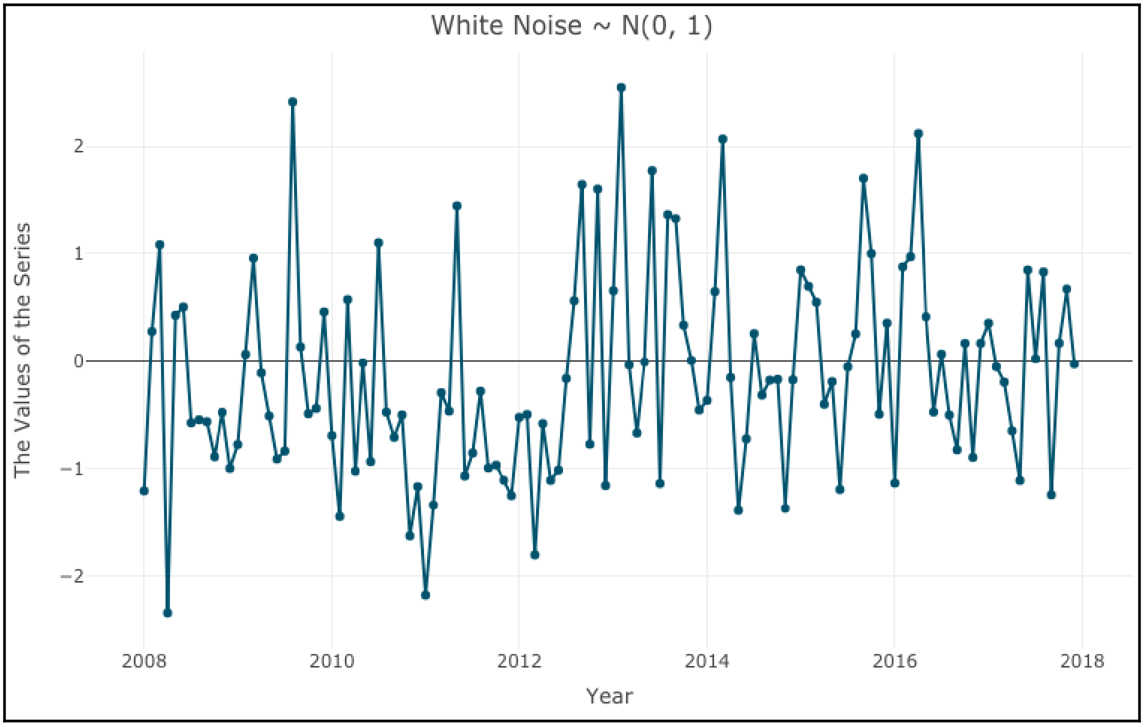

Different Types of Trends
Baseline_No_Trend
Positive_Linear_Trend
Negative_Linear_Trend
Exponential_Trend
The Values of the Series
11
10
9
8
7
6
5
4
3
2000
2002
2004
2006
2008
2010
2012
2014
2016
Year

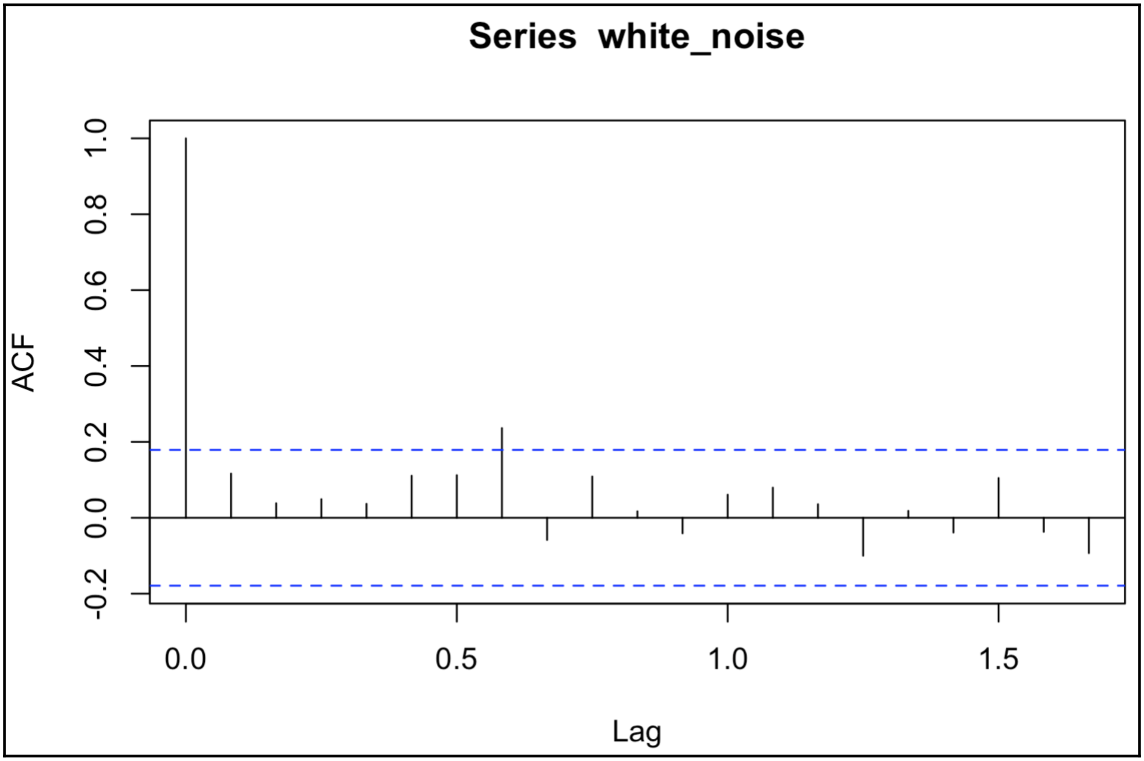

Seasonal Pattern without Trend
The Values of the Series
6
5.5
5
4.5
4
2000
2002
2004
2006
2008
2010
2012
2014
2016
Year

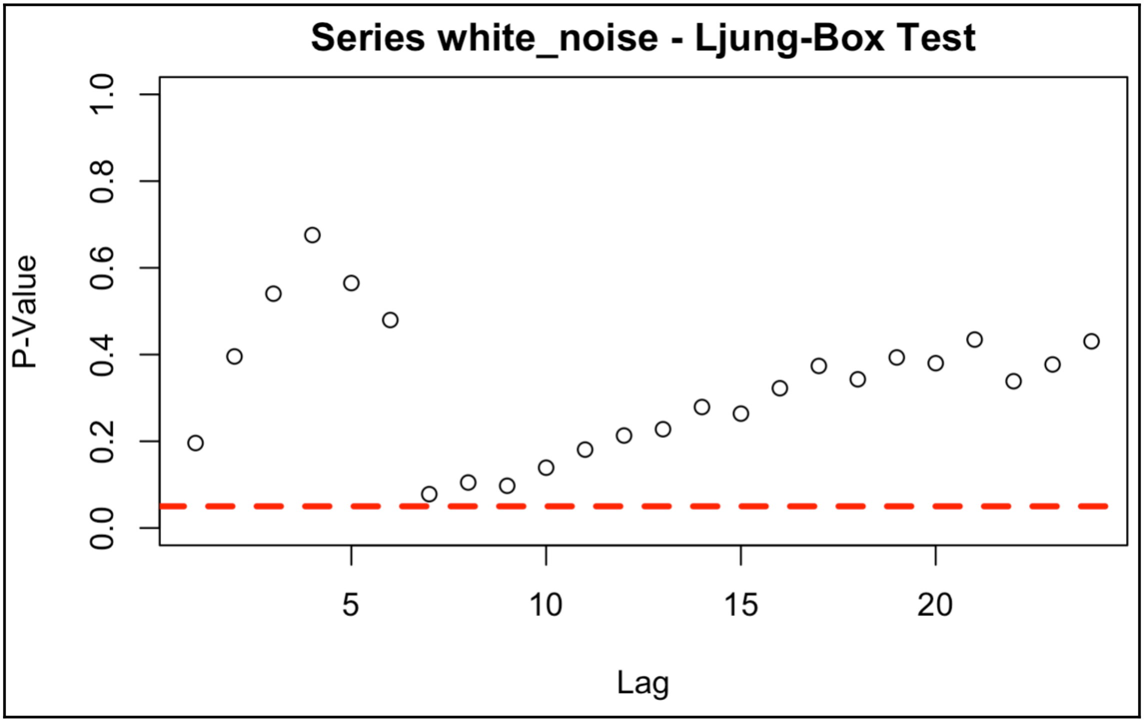

Seasonal Pattern with Trend
Positive_Linear_Trend
Negative_Linear_Trend
Exponential_Trend
The Values of the Series
12
10
8
6
4
2
2000
2002
2004
2006
2008
2010
2012
2014
2016
Year

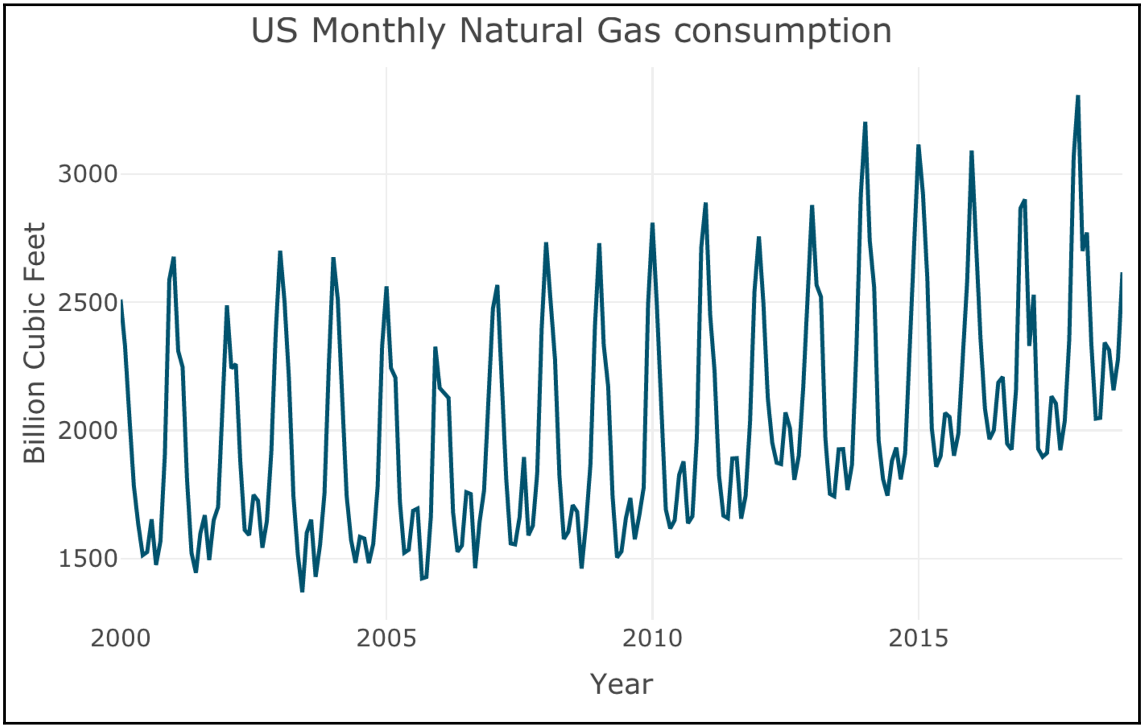

Heatmap - US Natural Gas Consumption
Dec
Nov
Oct
Sep
Aug
Jul
Jun
May
Apr
Mar
Feb
Jan
Month
2000
2005
2010
2015
Year
3000
2500
2000
1500

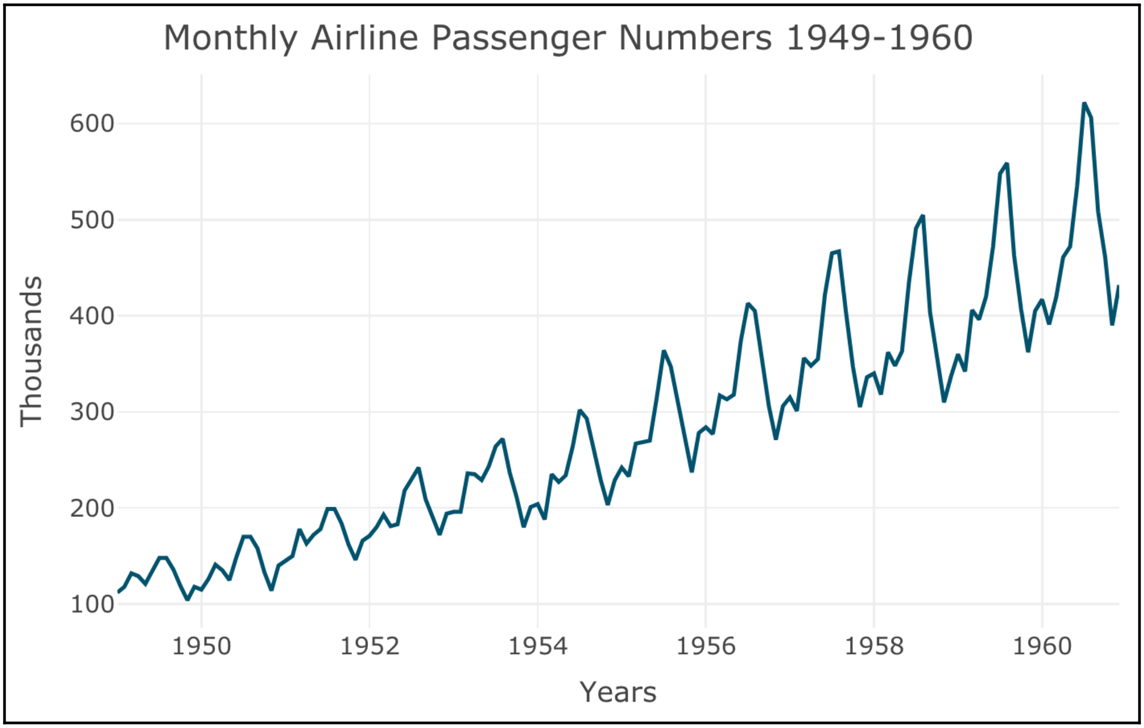

Heatmap - US Unemployment Rate
Dec
Nov
Oct
Sep
Aug
Jul
Jun
May
Apr
Mar
Feb
Jan
Month
1950
1960
1970
1980
1990
2000
2010
Year
10
8
6
4

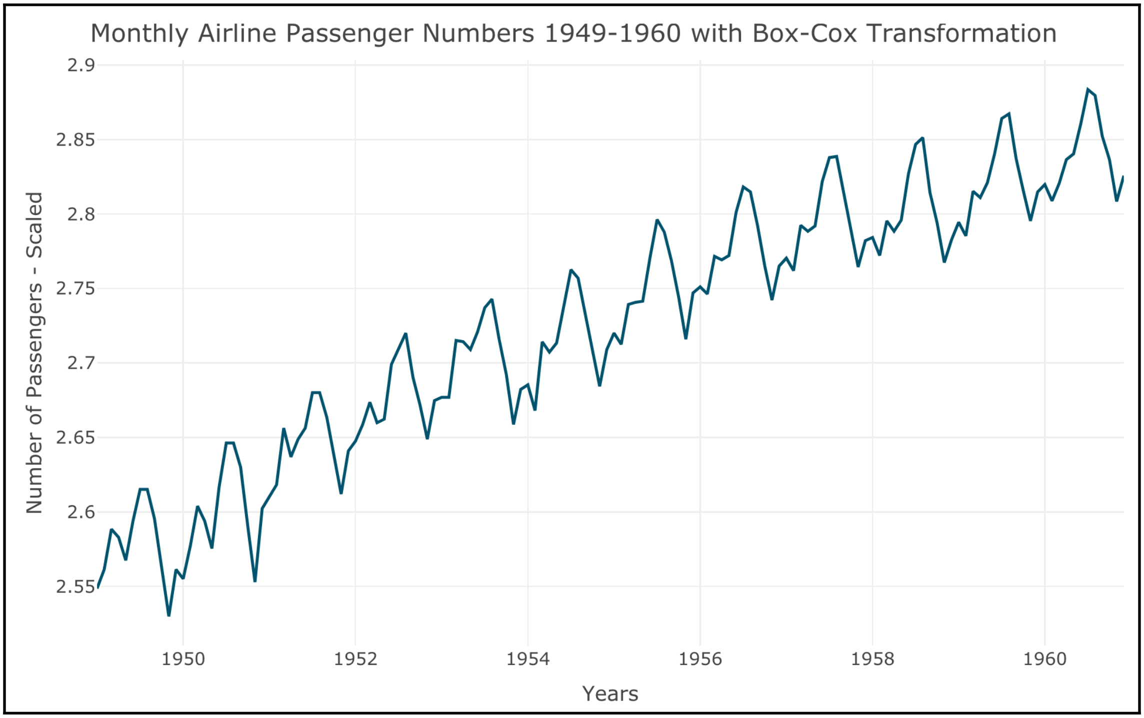

White Noise ~ N(0, 1)
The Values of the Series
2
1
0
−1
−2
2008
2010
2012
2014
2016
2018
Year

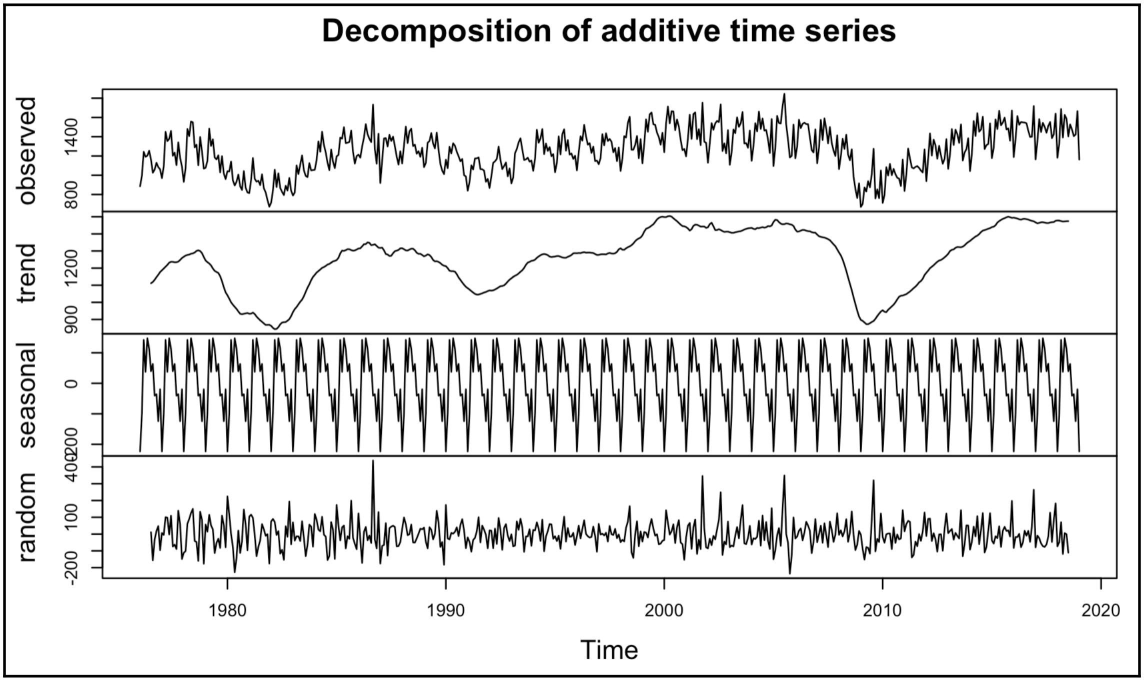

Series white_noise
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0.0
0.5
1.0
1.5
Lag

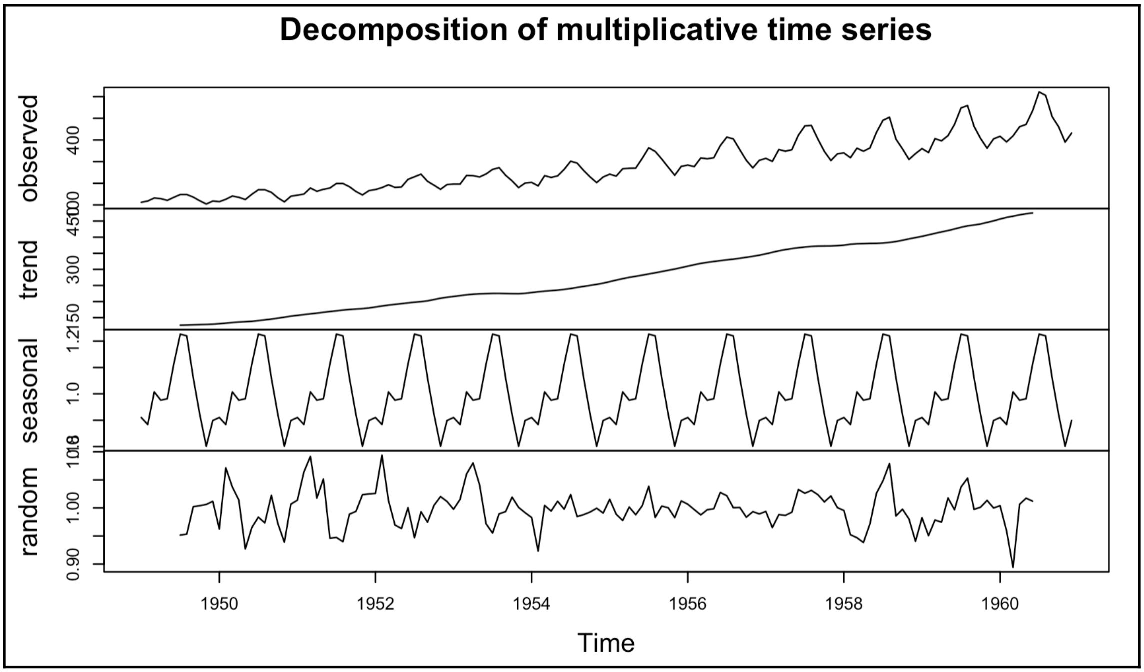

Series white_noise - Ljung-Box Test
P-Value
0.0
0.2
0.4
0.6
0.8
1.0
5
10
15
20
Lag


US Monthly Natural Gas consumption
Billion Cubic Feet
3000
2500
2000
1500
2000
2005
2010
2015
Year

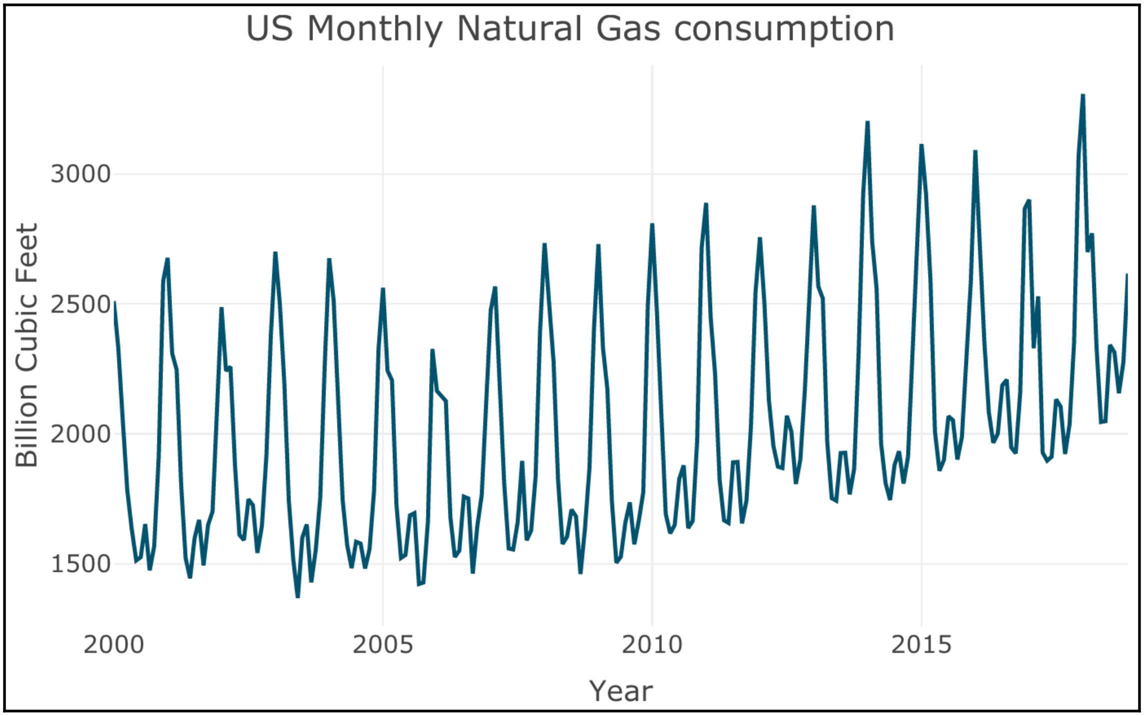

Monthly Airline Passenger Numbers 1949-1960
600
500
400
300
200
100
Thousands
1950
1952
1954
1956
1958
1960
Years

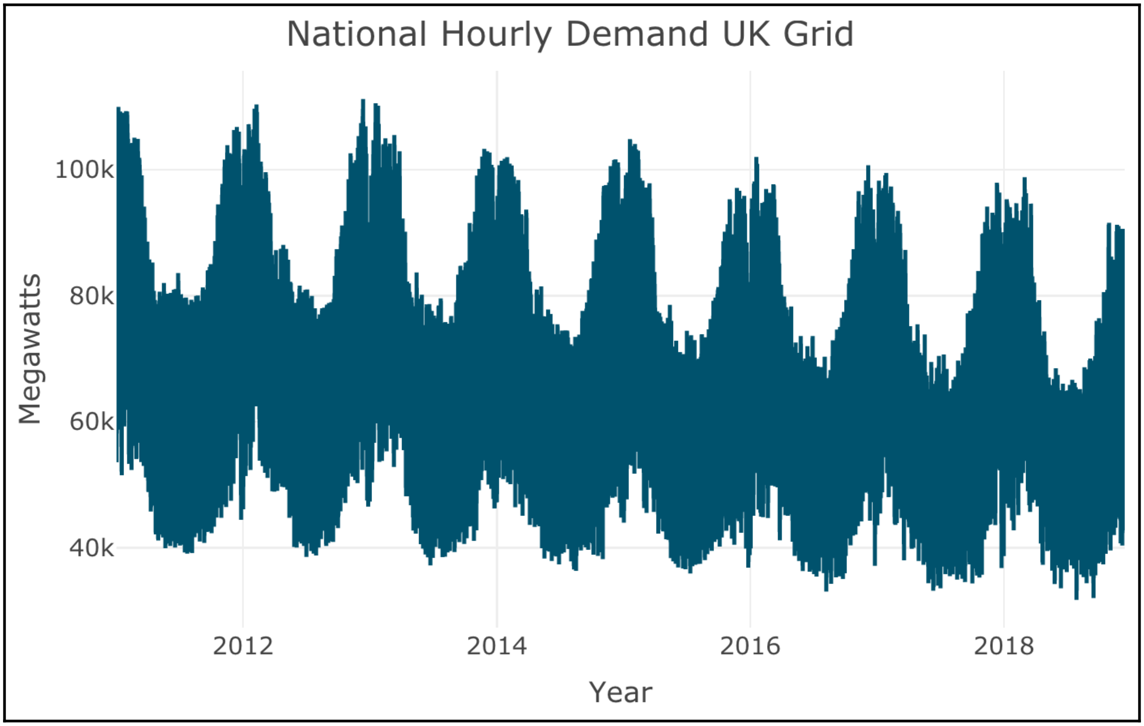

Monthly Airline Passenger Numbers 1949-1960 with Box-Cox Transformation
Number of Passengers - Scaled
Years
2.9
2.85
2.8
2.75
2.7
2.65
2.6
2.55
1950
1952
1954
1956
1958
1960

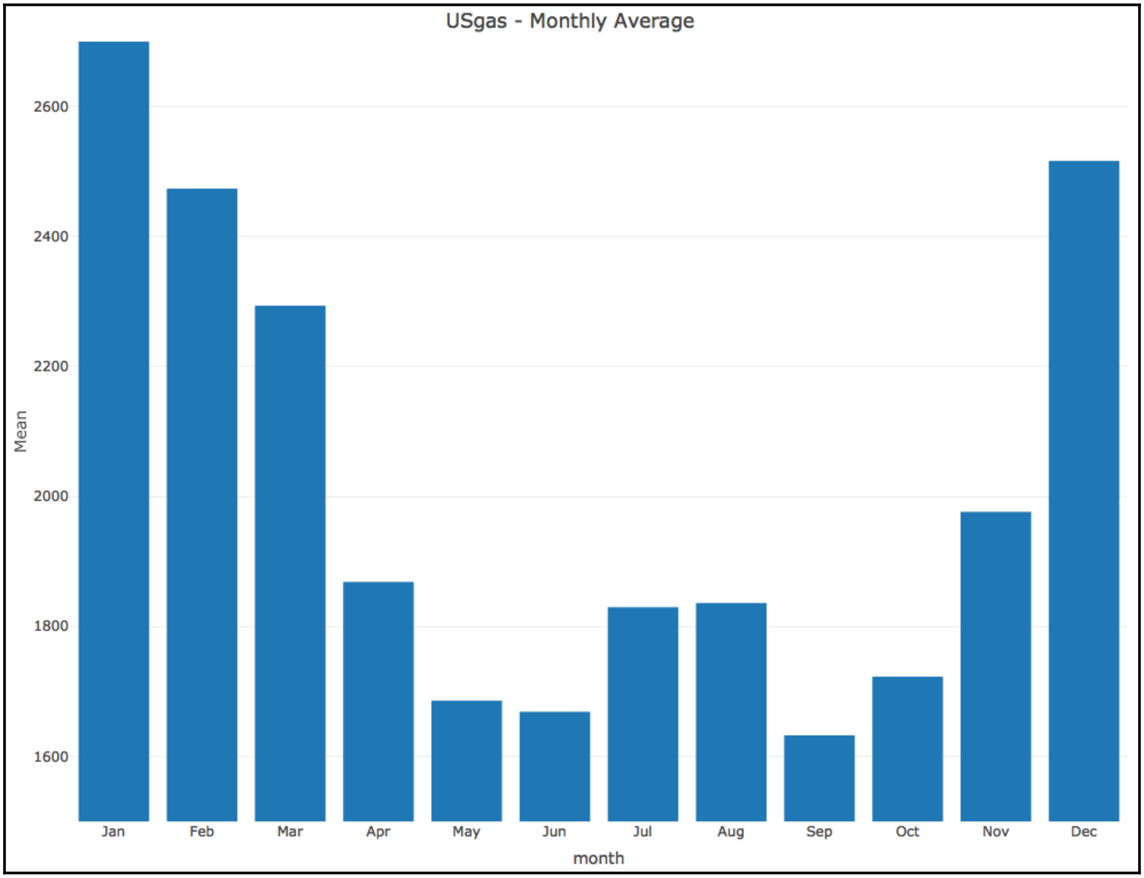

Decomposition of additive time series
observed
trend
seasonal
random
800
1400
900
1200
0
-200
400
100
-200
1980
1990
2000
2010
2020
Time

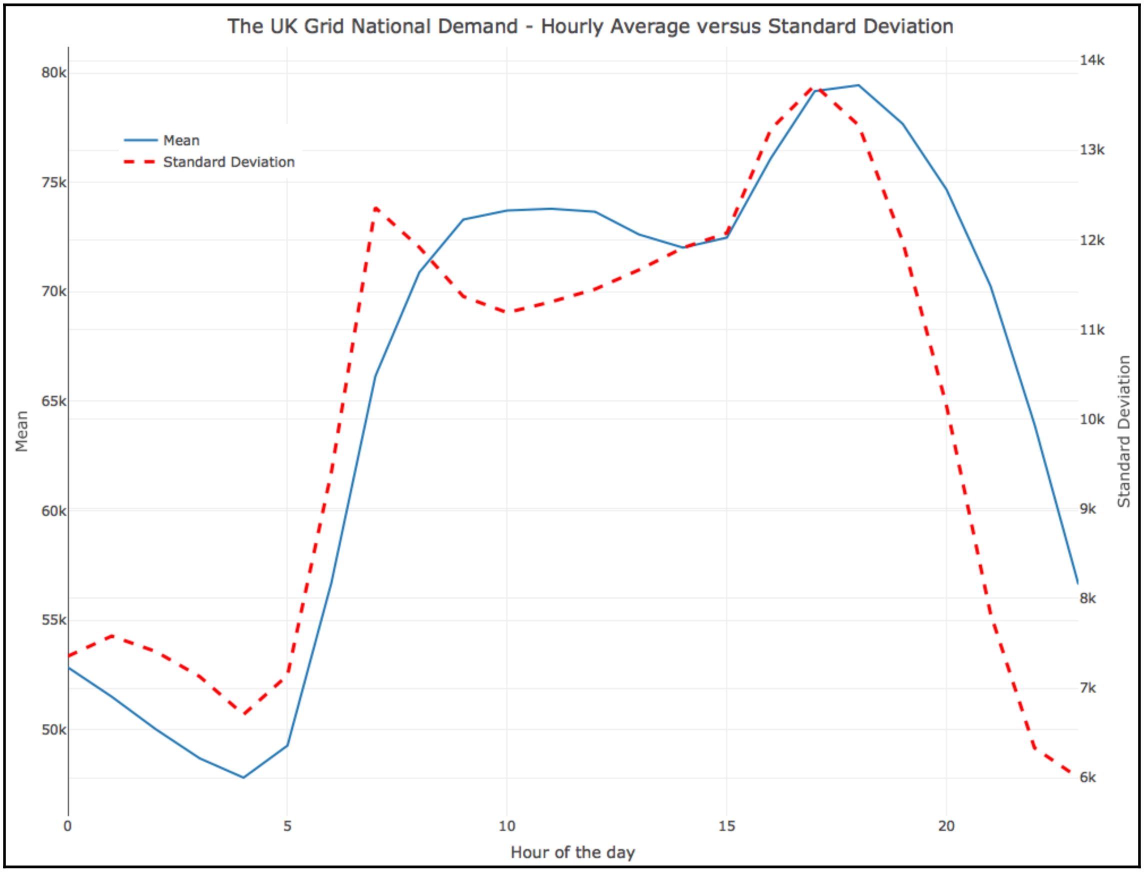

Decomposition of multiplicative time series
observed
trend
seasonal
random
400
150
300
450
1.0
1.2
0.90
1.00
1950
1952
1954
1956
1958
1960
Time

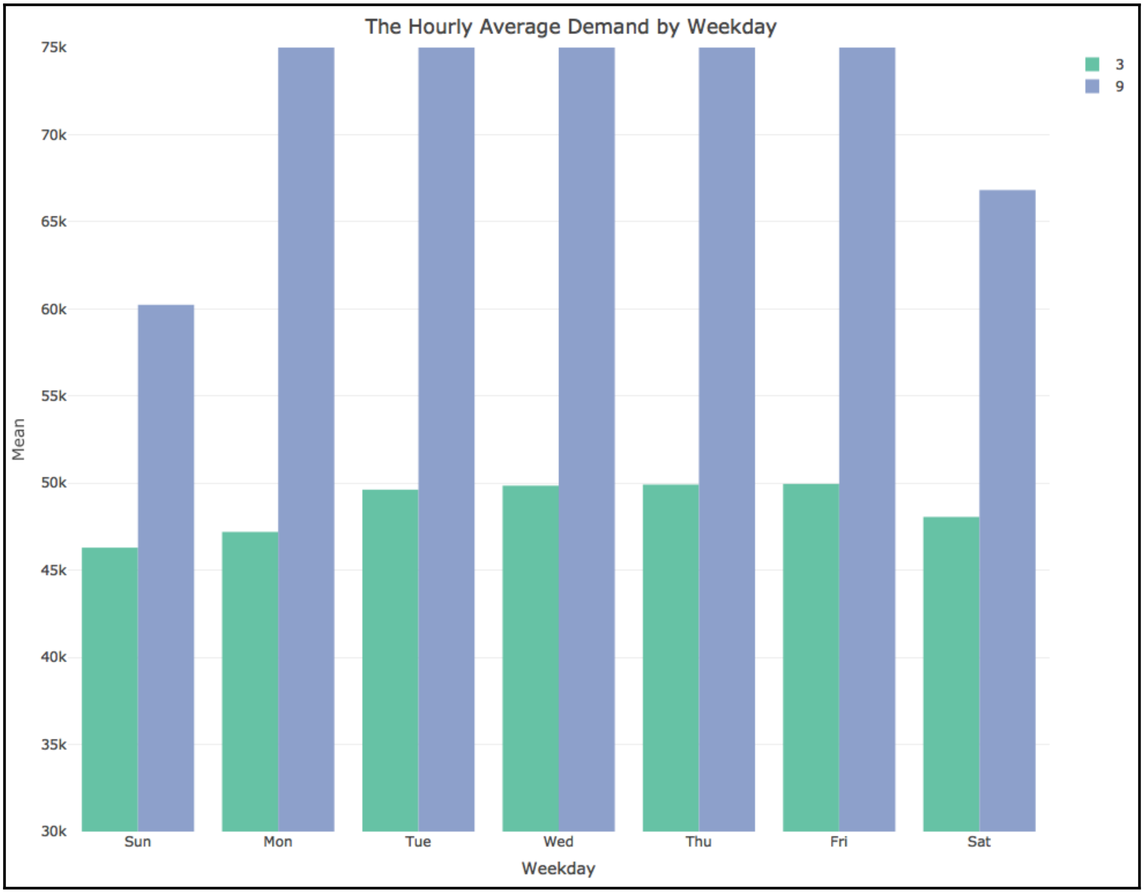
Chapter 6: Seasonality Analysis

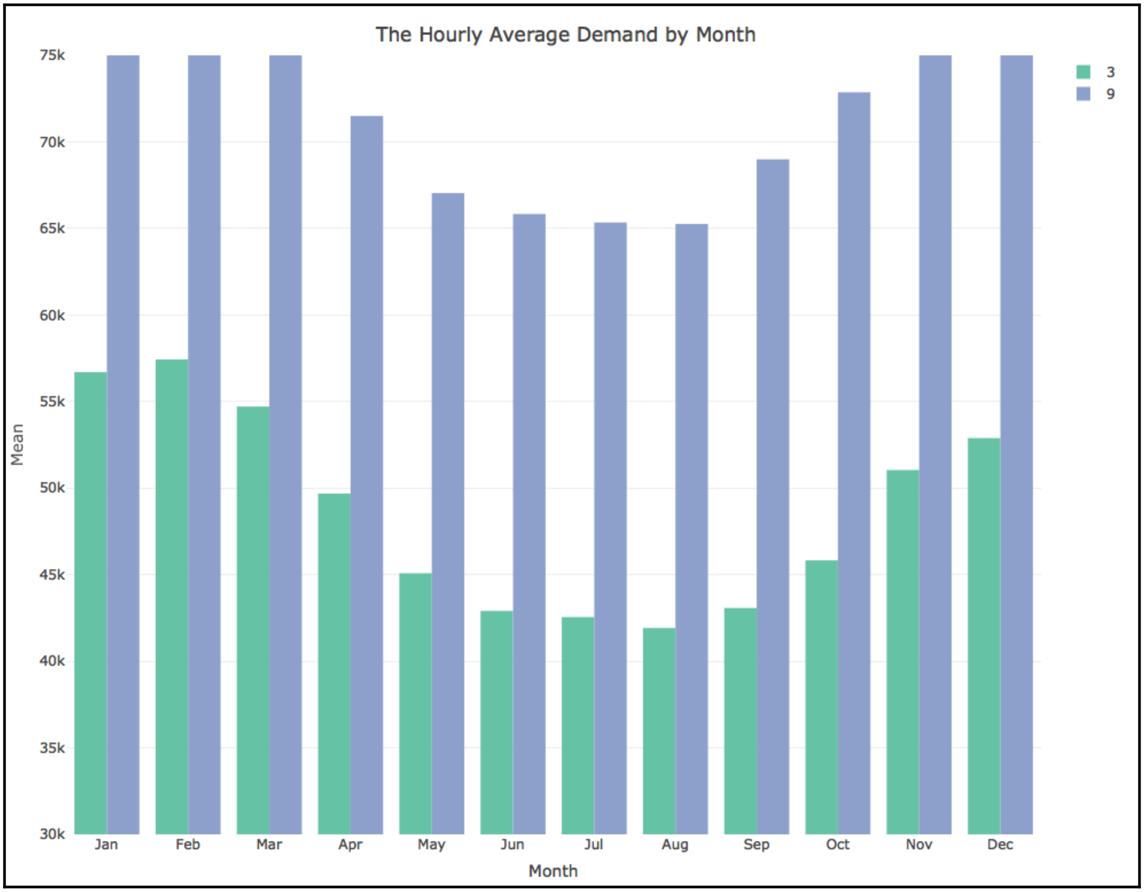


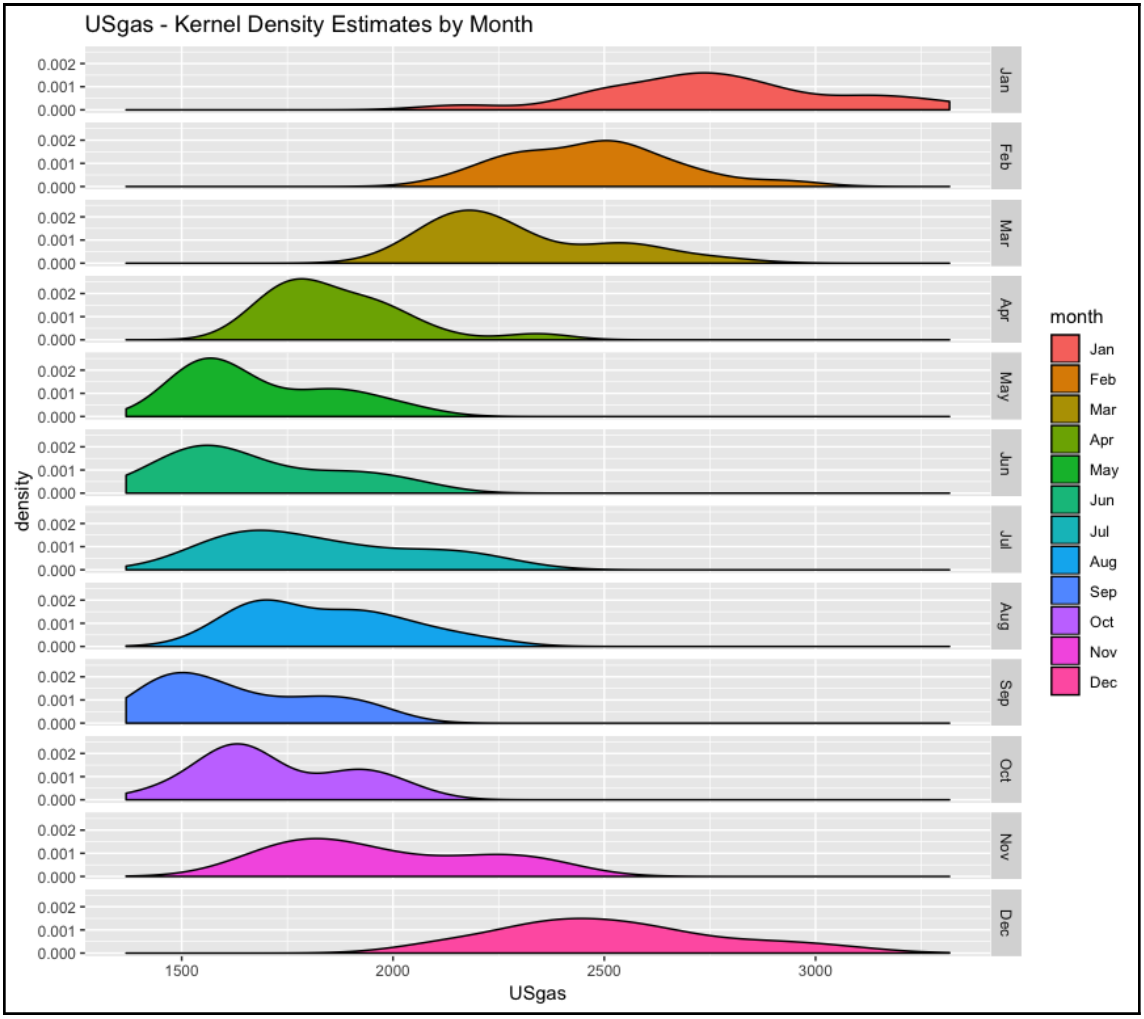

National Hourly Demand UK Grid
Megawatts
100k
80k
60k
40k
2012
2014
2016
2018
Year


USgas - Monthly Average
2600
2400
2200
2000
1800
1600
Mean
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
month

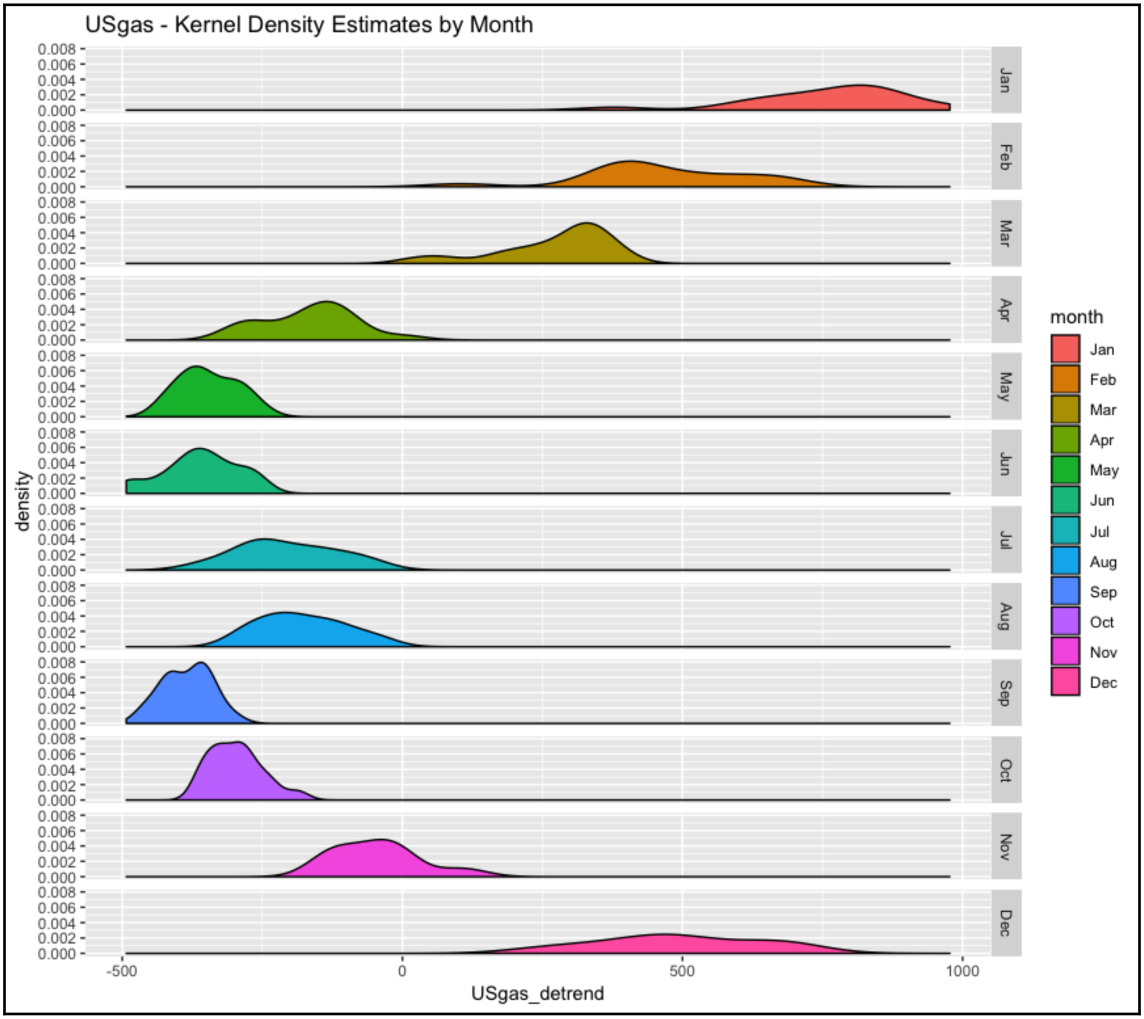

The UK Grid National Demand - Hourly Average versus Standard Deviation
Mean
Standard Deviation
80k
75k
70k
65k
60k
55k
50k
Mean
14k
13k
12k
11k
10k
9k
8k
7k
6k
Standard Deviation
0
5
10
15
20
Hour of the day

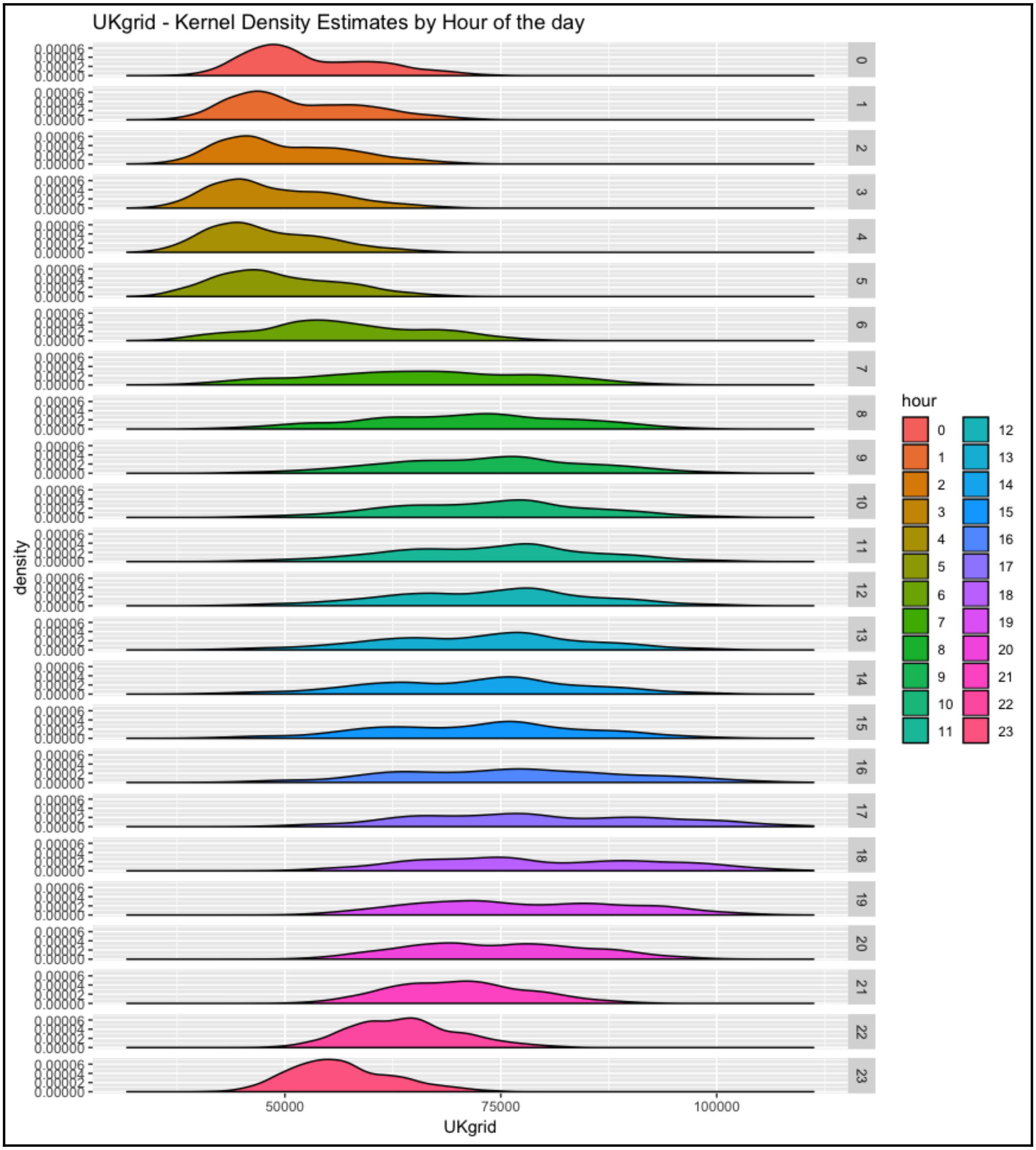

The Hourly Average Demand by Weekday
75k
70k
65k
60k
55k
50k
45k
40k
35k
30k
Mean
Sun
Mon
Tue
Wed
Thu
Fri
Sat
Weekday
3
9

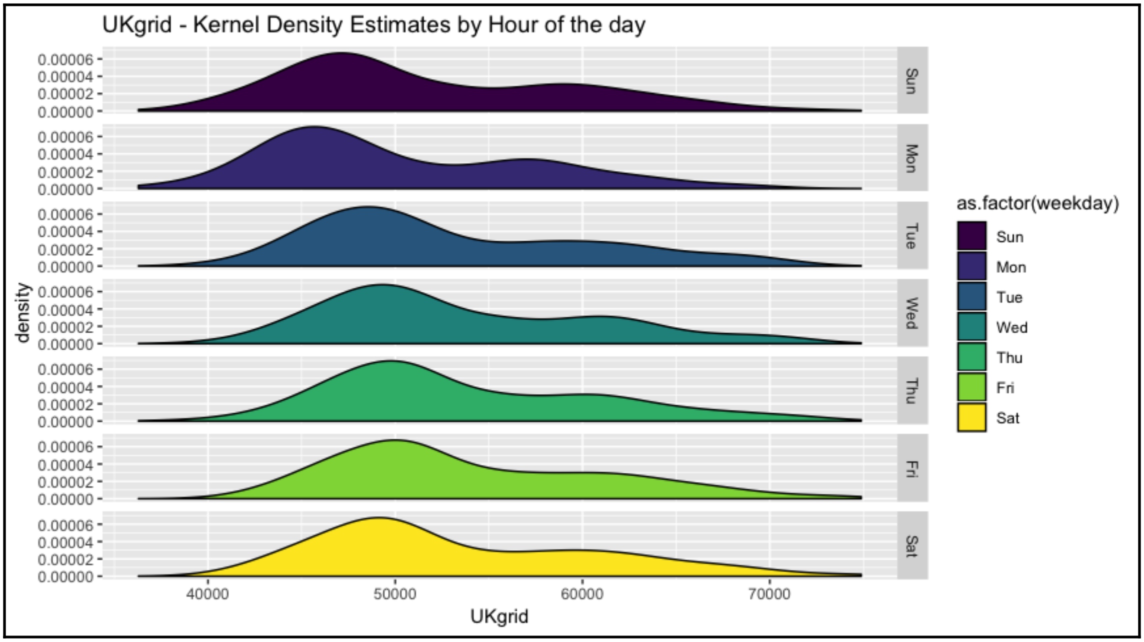

The Hourly Average Demand by Month
75k
70k
65k
60k
55k
50k
45k
40k
35k
30k
Mean
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
3
9

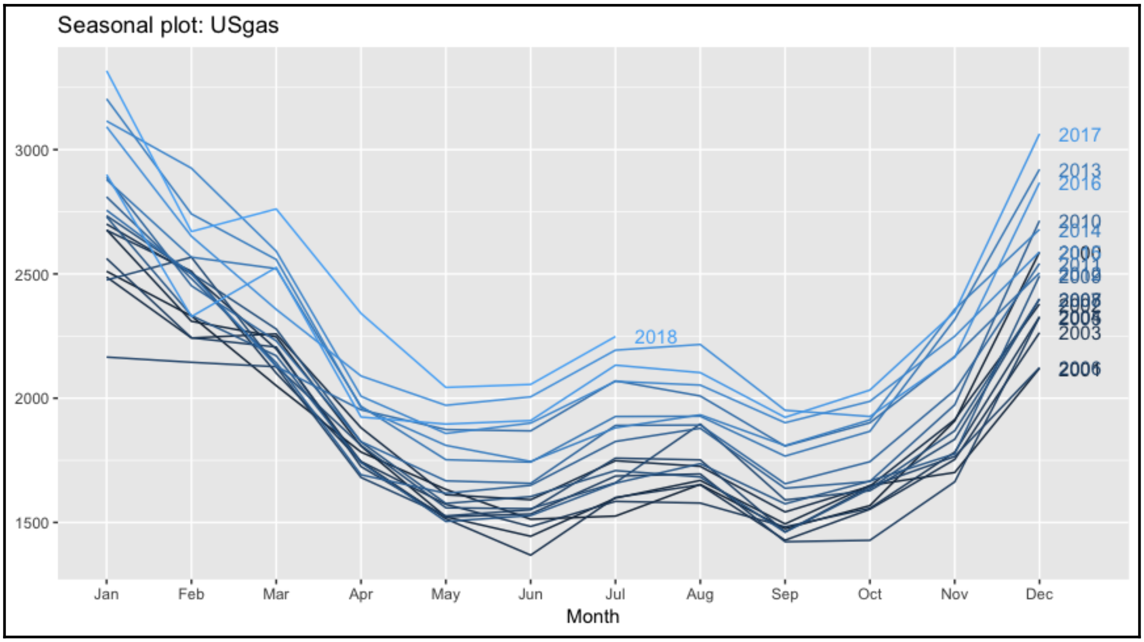


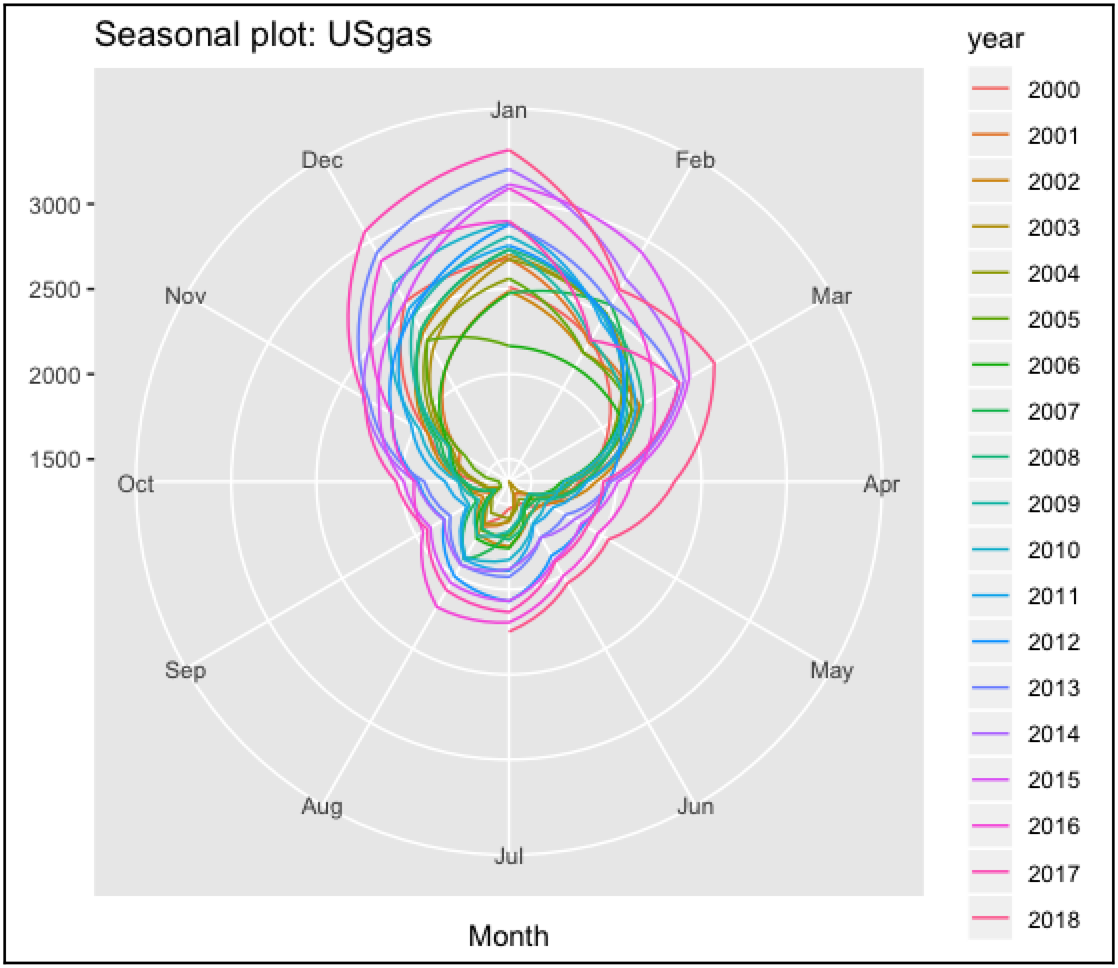
USgas - Kernel Density Estimates by Month
density
0.002
0.001
0.000
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
1500
2000
2500
3000
USgas
month

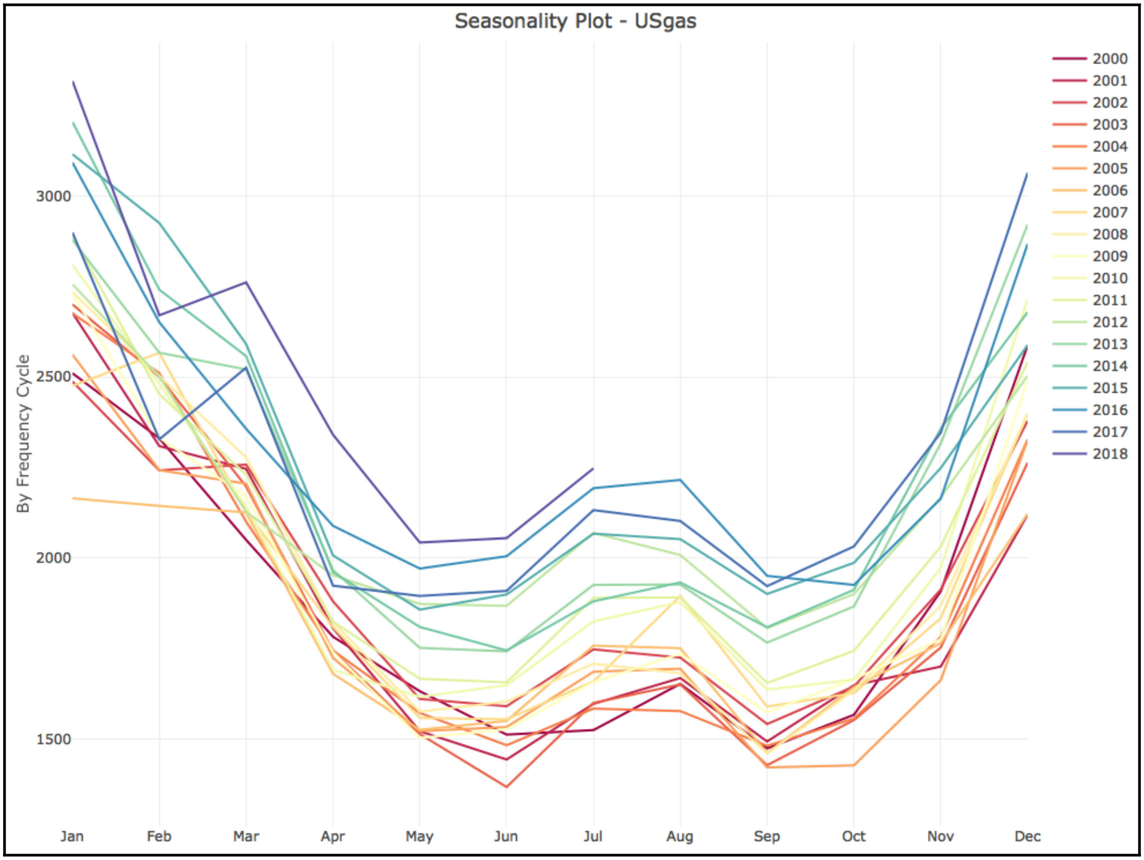

USgas - Kernel Density Estimates by Month
density
USgas_detrend
-500
0
500
1000
0.000
0.002
0.004
0.006
0.008
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
month


UKgrid - Kernel Density Estimates by Hour of the day
density
UKgrid
50000
75000
100000
hour
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

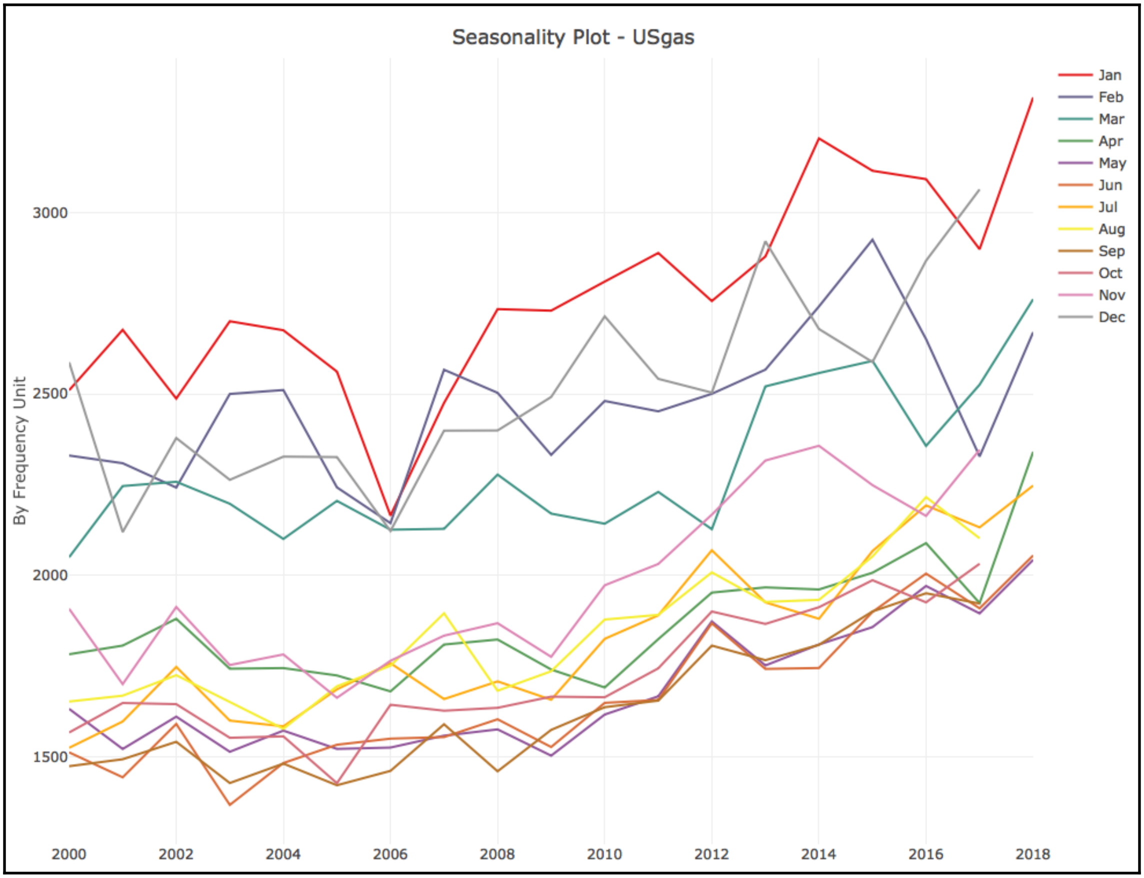

UKgrid - Kernel Density Estimates by Hour of the day
0.00006
0.00004
0.00002
0.00000
density
Sun
Mon
Tue
Wed
Thu
Fri
Sat
40000
50000
60000
70000
UKgrid
as.factor(weekday)

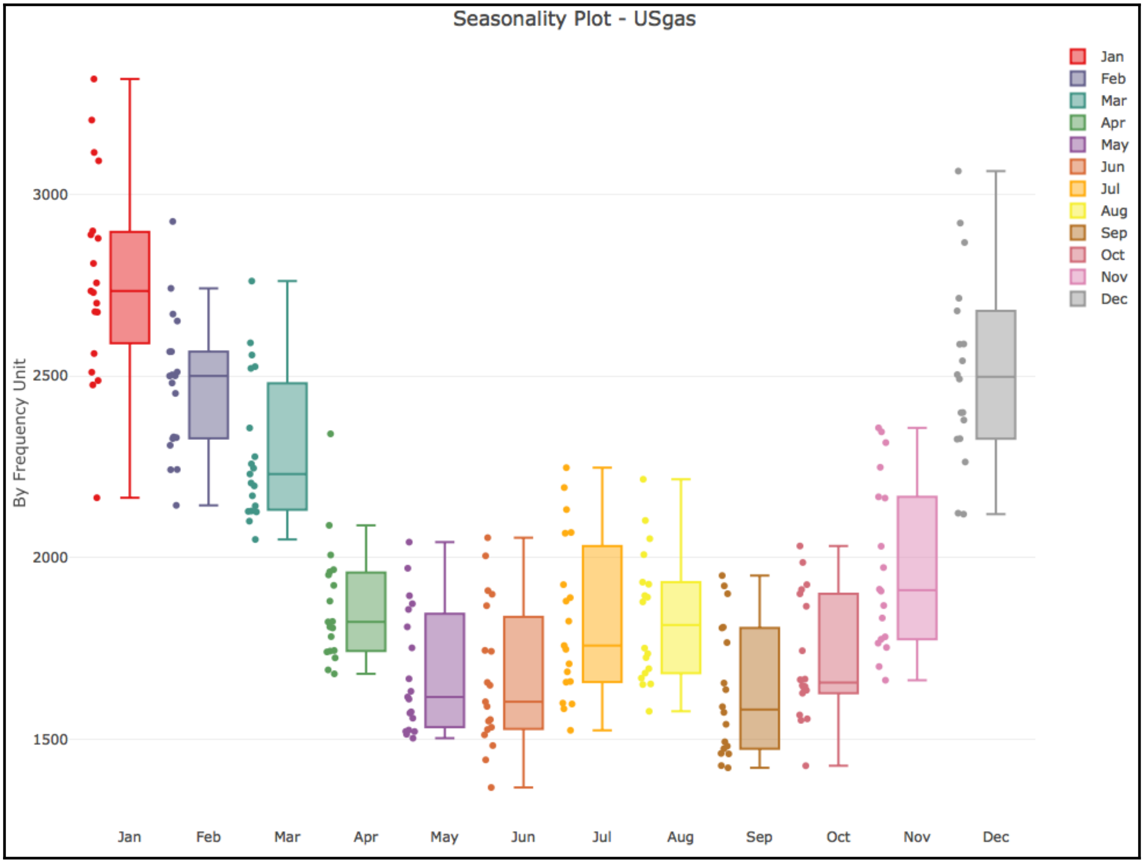

Seasonal plot: USgas
3000
2500
2000
1500
2017
2018
2003
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

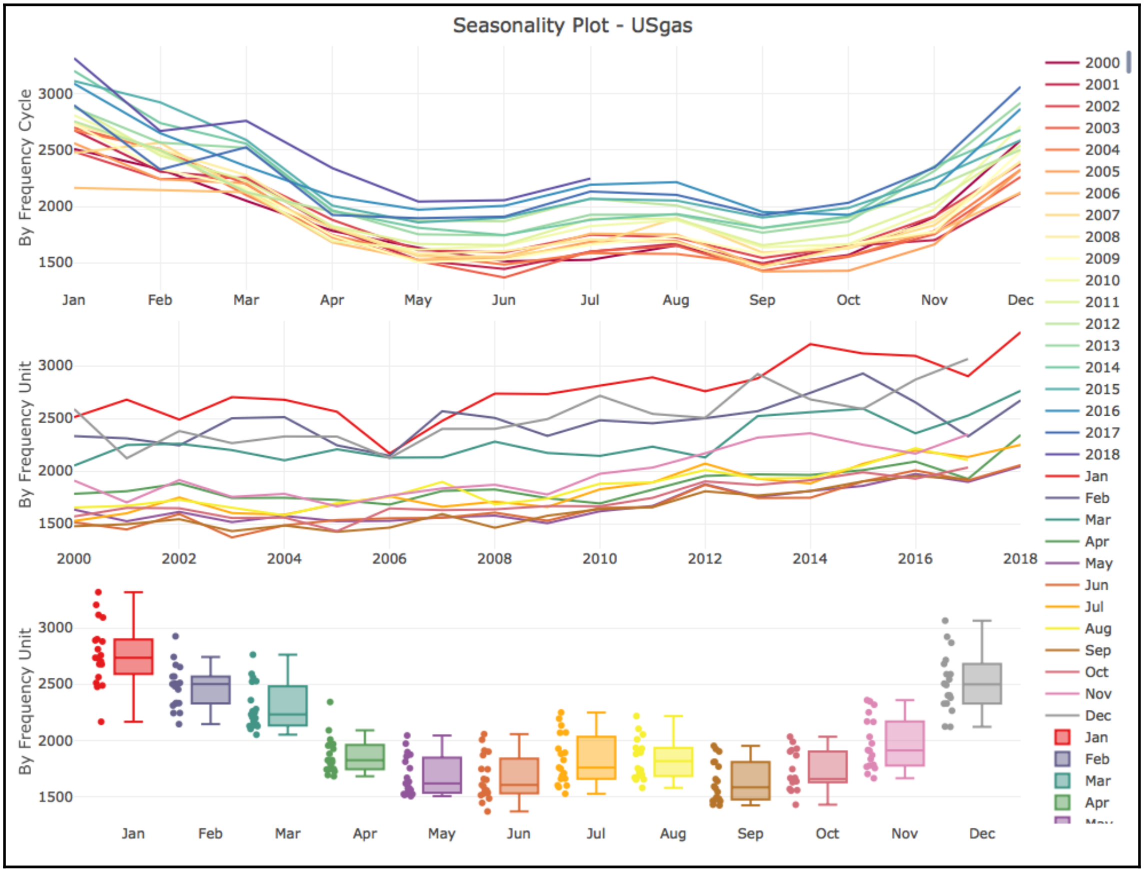

Seasonal plot: USgas
year
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
3000
2500
2000
1500
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

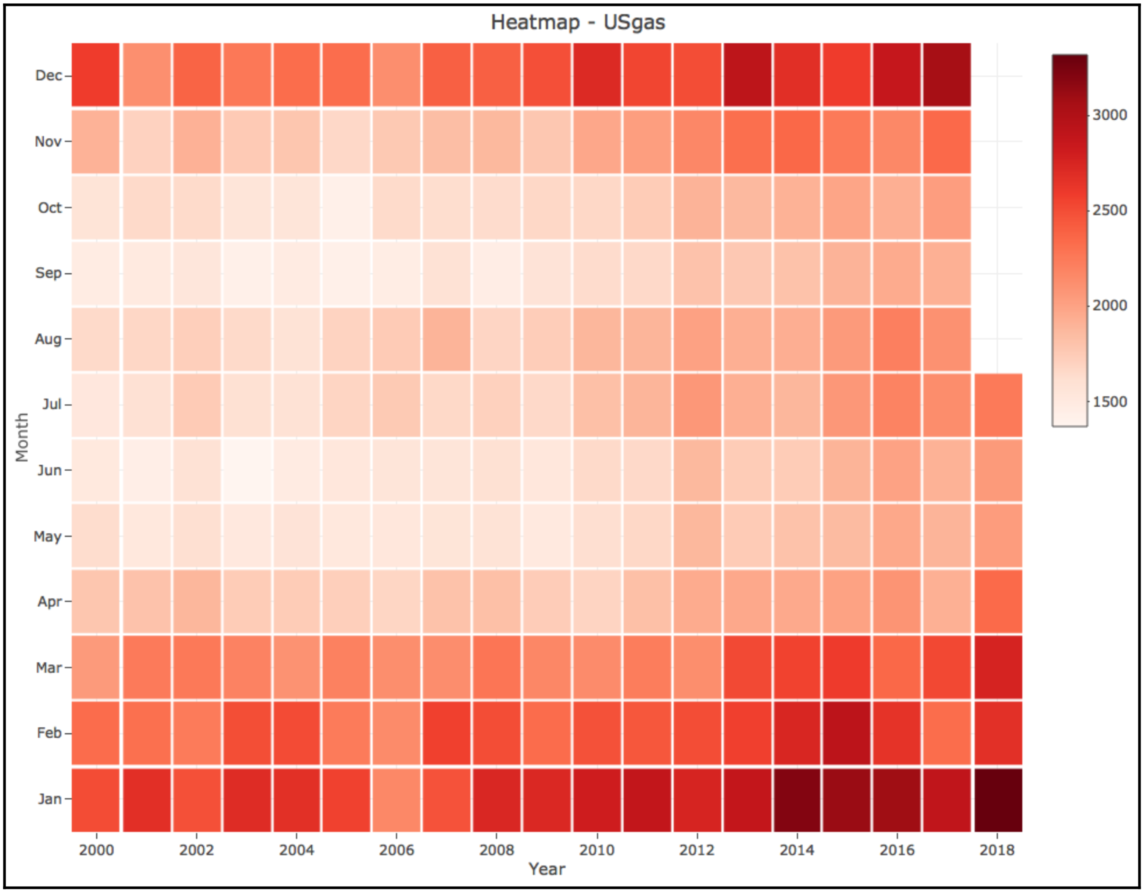


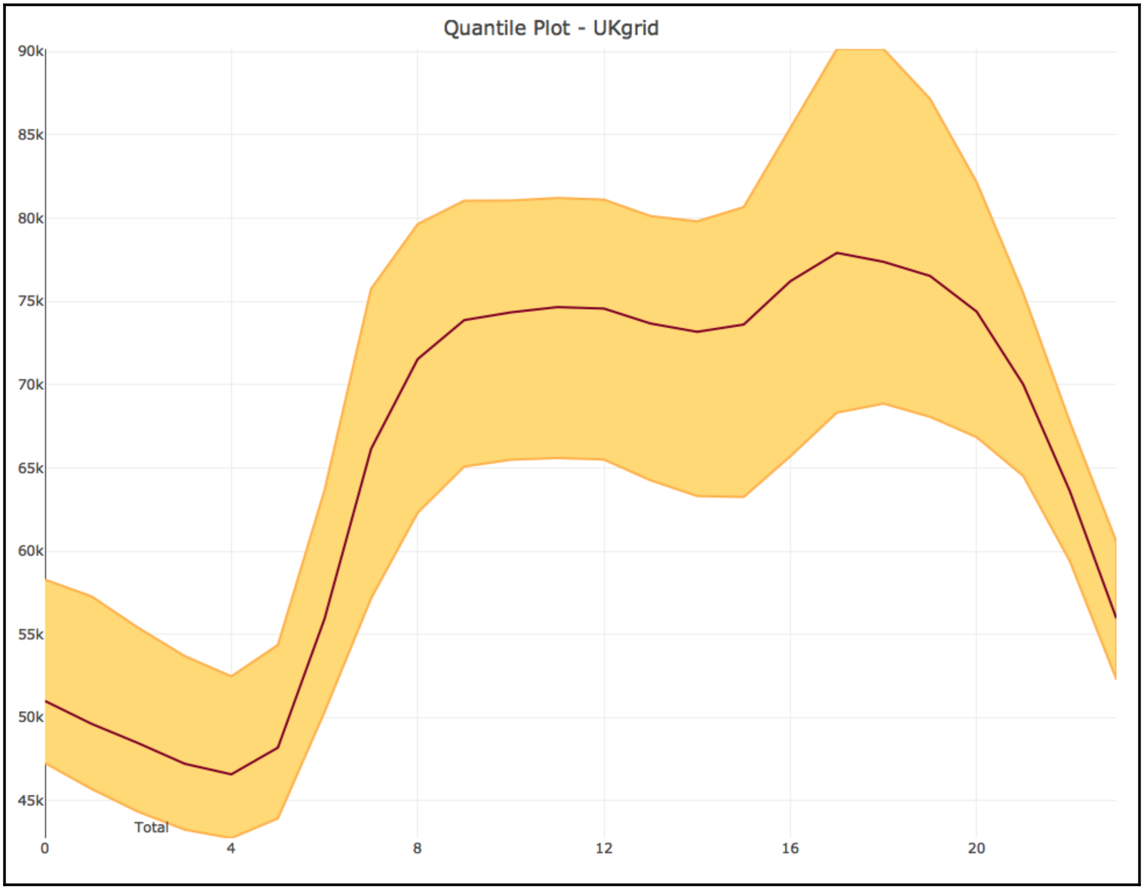
Seasonality Plot - USgas
By Frequency Cycle
3000
2500
2000
1500
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

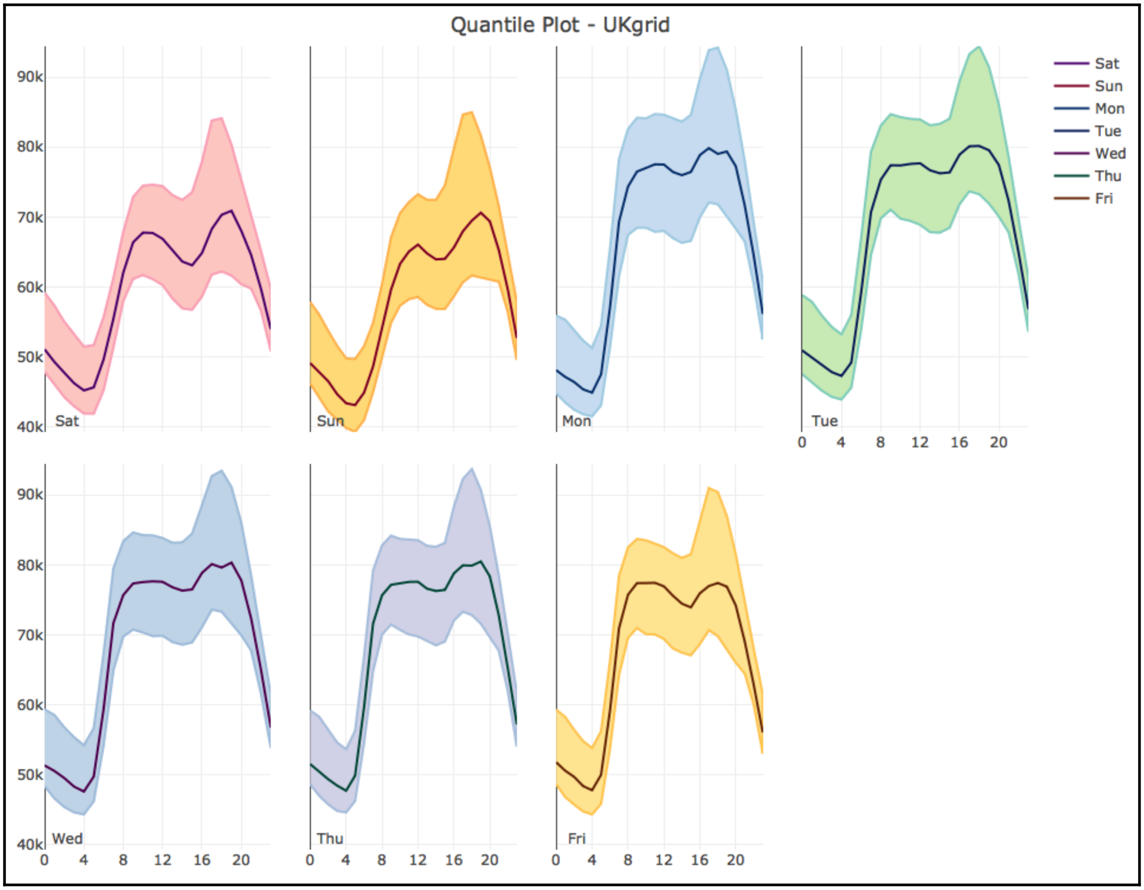

Seasonality Plot - USgas
By Frequency Unit
3000
2500
2000
1500
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

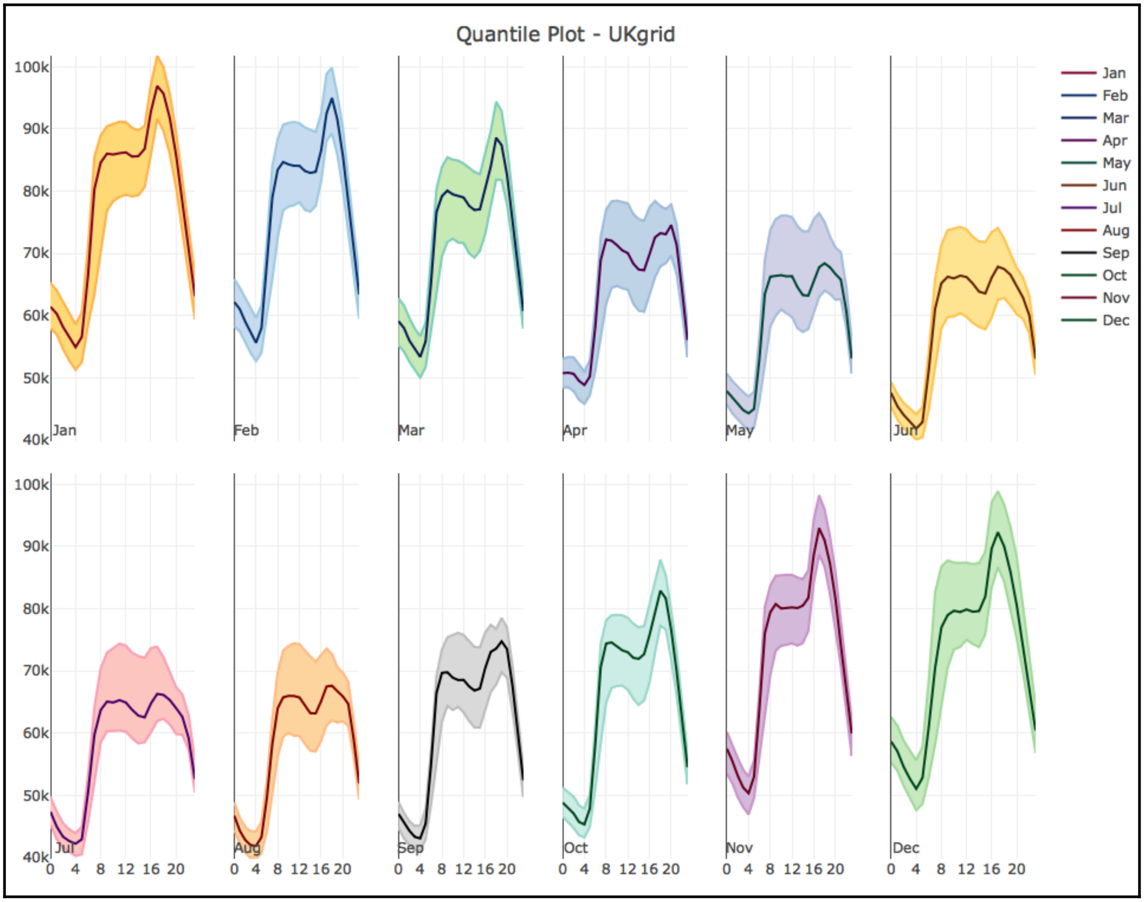

Seasonality Plot - USgas
By Frequency Unit
3000
2500
2000
1500
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

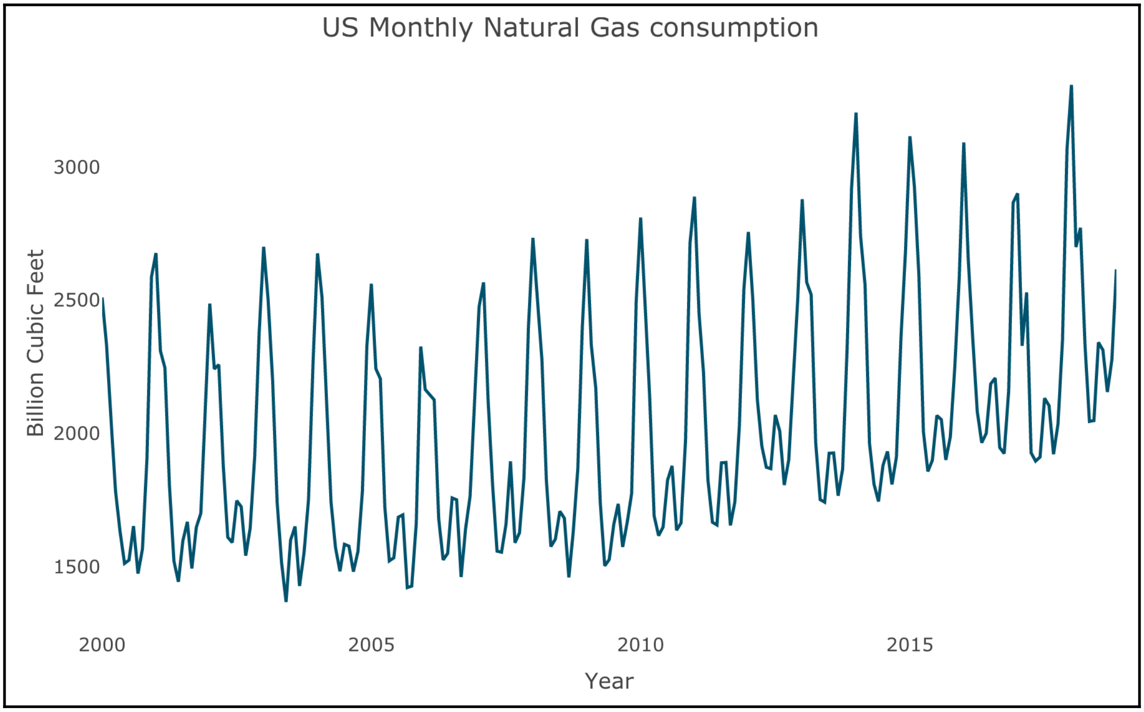

Seasonality Plot - USgas
By Frequency Cycle
By Frequency Unit
By Frequency Unit
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

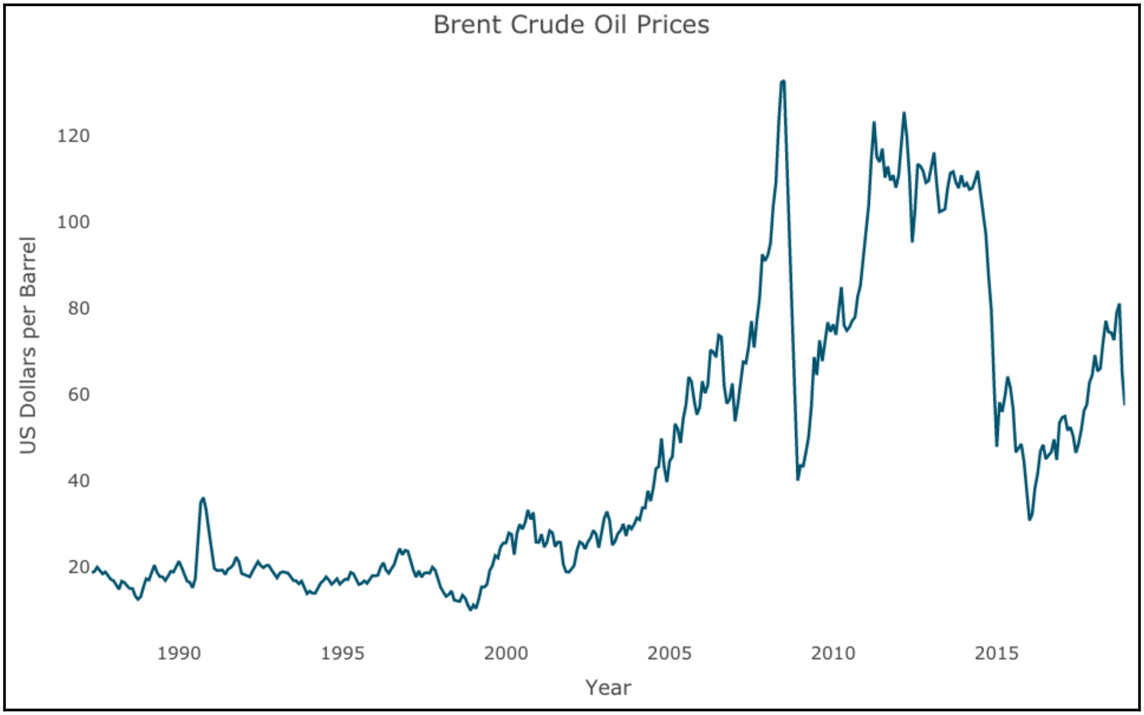

Heatmap - USgas
Month
Dec
Nov
Oct
Sep
Aug
Jul
Jun
May
Apr
Mar
Feb
Jan
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year
3000
2500
2000
1500

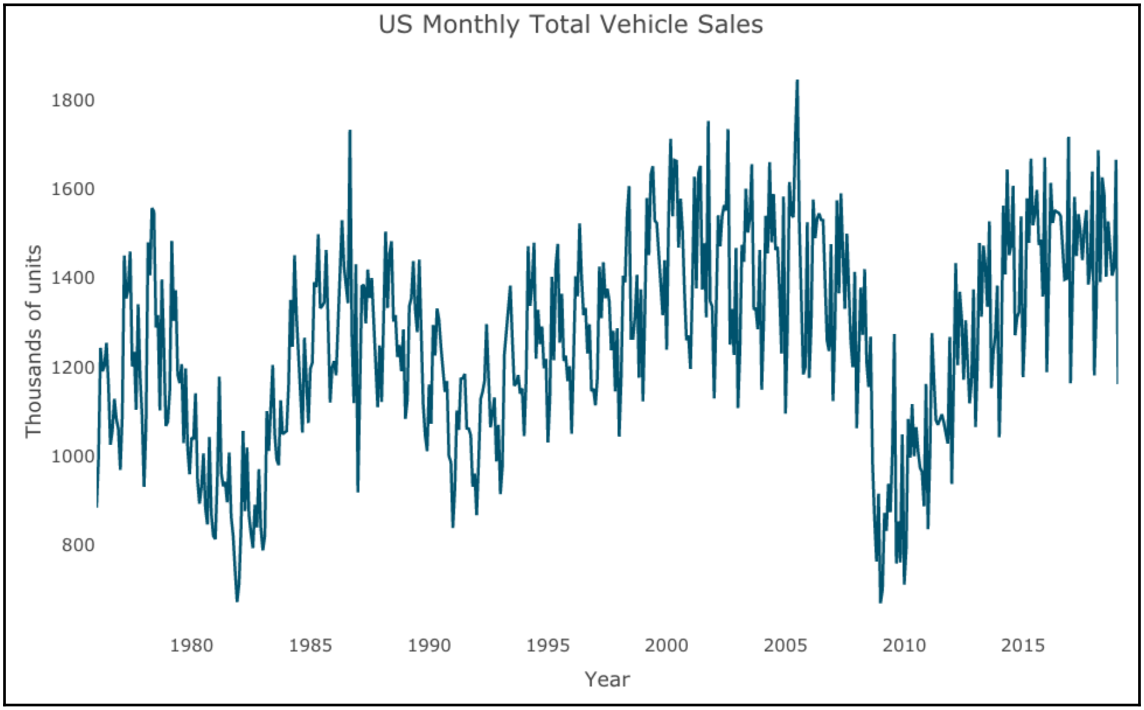

Quantile Plot - UKgrid
90k
85k
80k
75k
70k
65k
60k
55k
50k
45k
Total
0
4
8
12
16
20

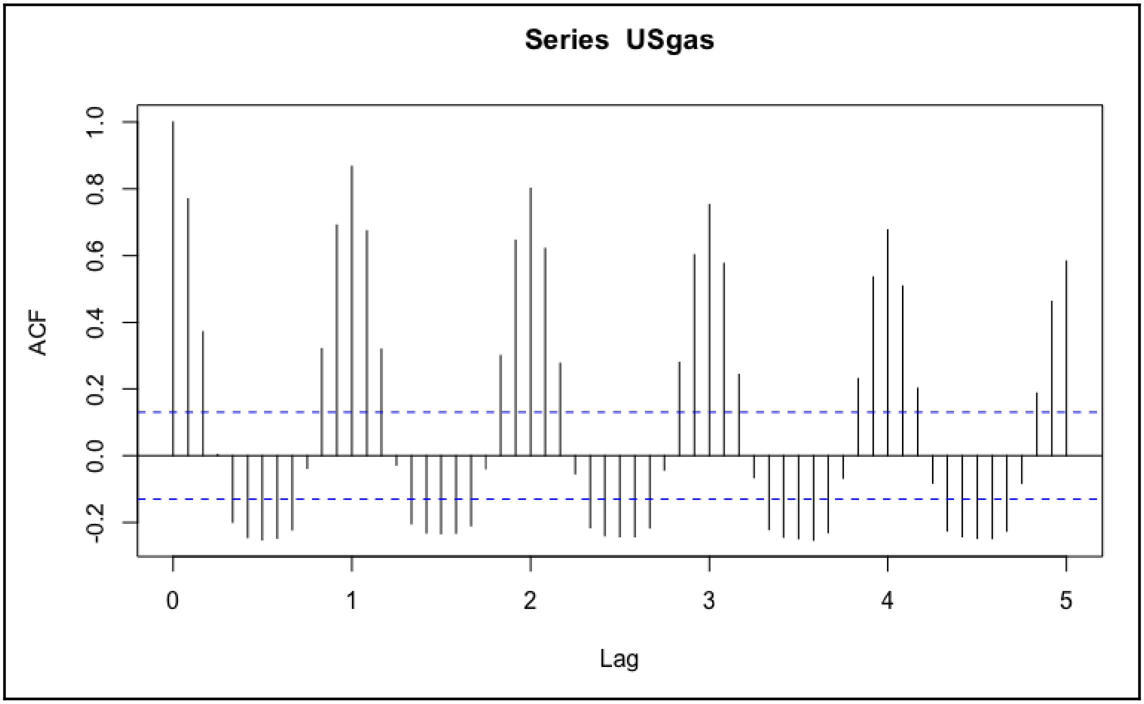

Quantile Plot - UKgrid
Sat
Sun
Mon
Tue
Wed
Thu
Fri
90k
80k
70k
60k
50k
40k
0
4
8
12
16
20

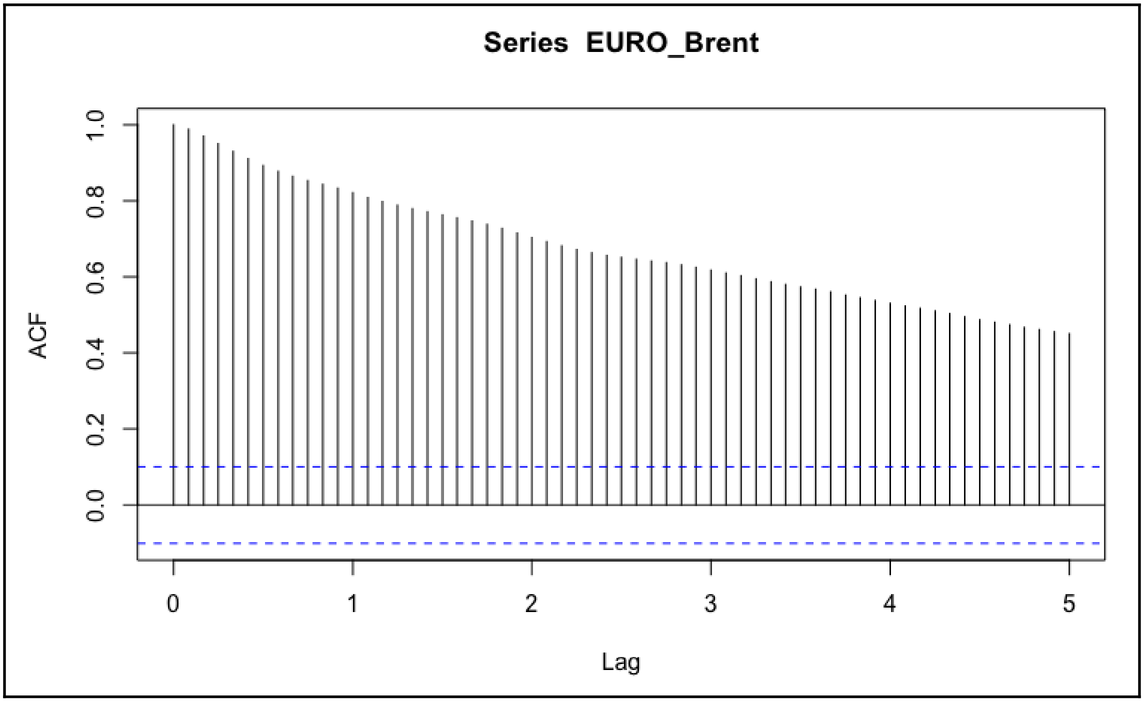

Quantile Plot - UKgrid
100k
90k
80k
70k
60k
50k
40k
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
0 4 8 12 16 20

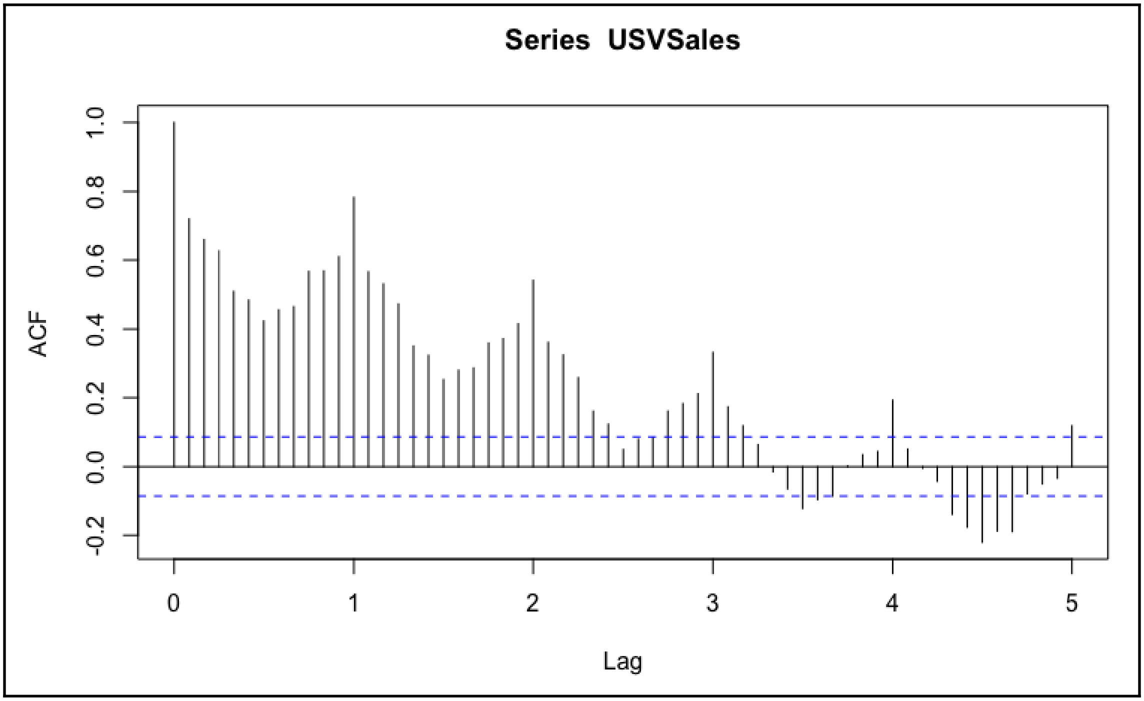

US Monthly Natural Gas consumption
Billion Cubic Feet
3000
2500
2000
1500
2000
2005
2010
2015
Year

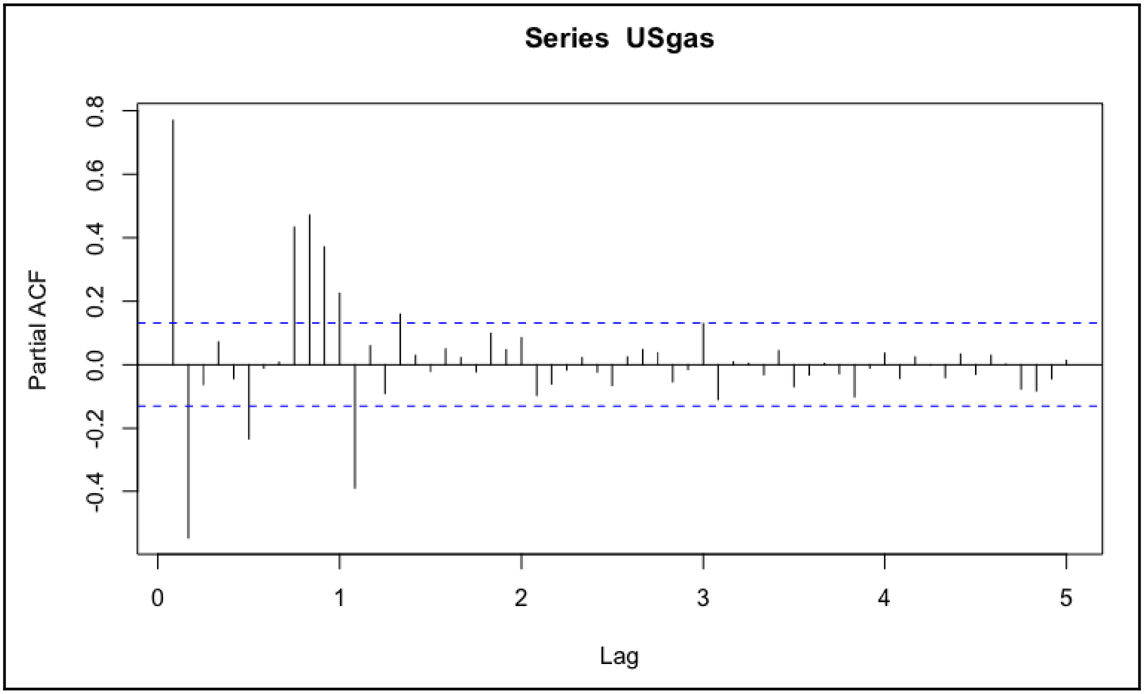

Brent Crude Oil Prices
US Dollars per Barrel
120
100
80
60
40
20
1990
1995
2000
2005
2010
2015
Year

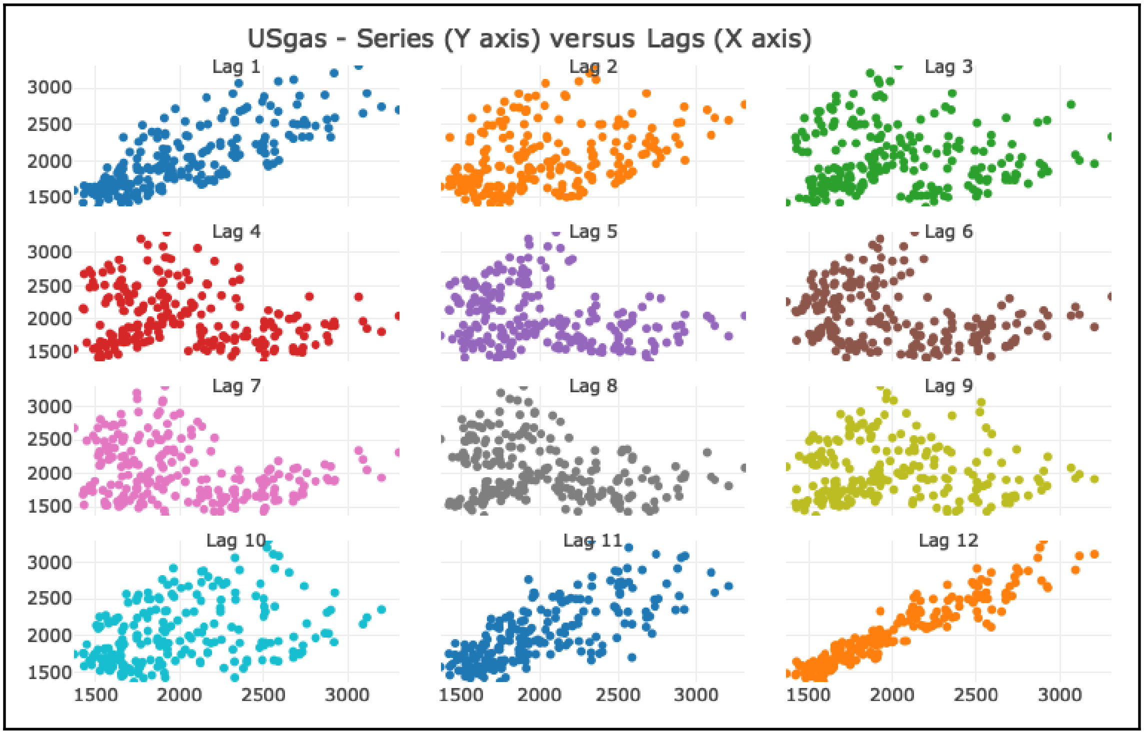

US Monthly Total Vehicle Sales
1800
1600
1400
1200
1000
800
Thousands of units
1980
1985
1990
1995
2000
2005
2010
2015
Year

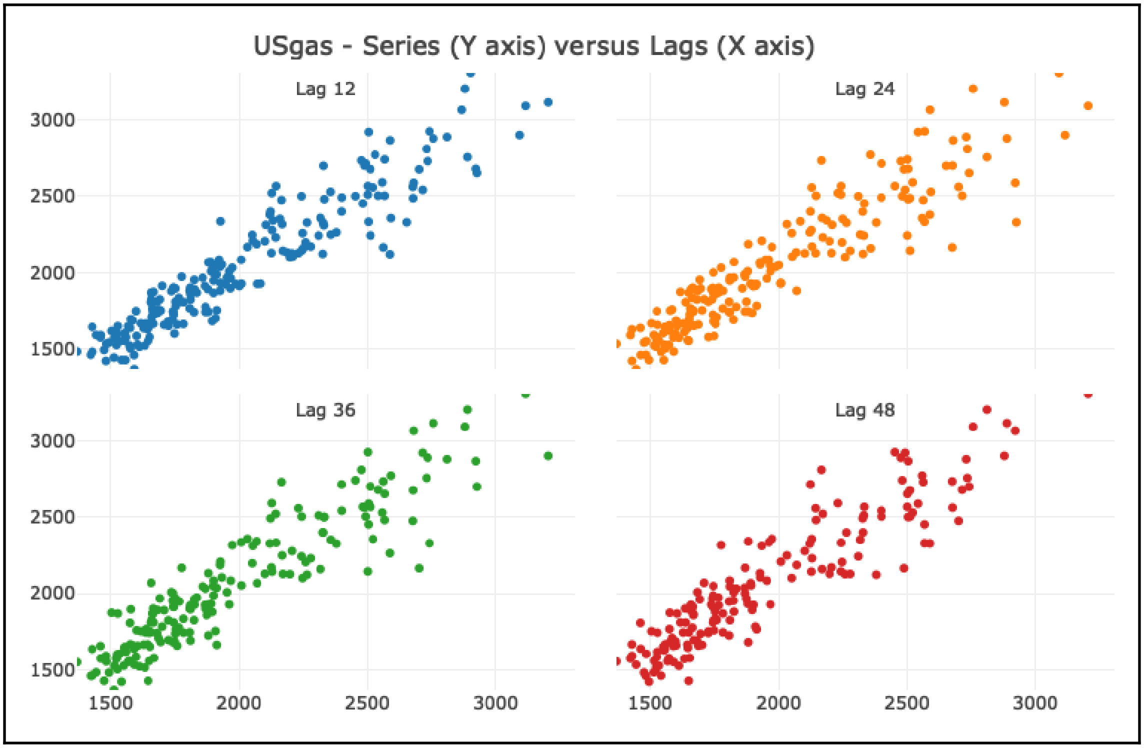

Series USgas
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
1
2
3
4
5
Lag

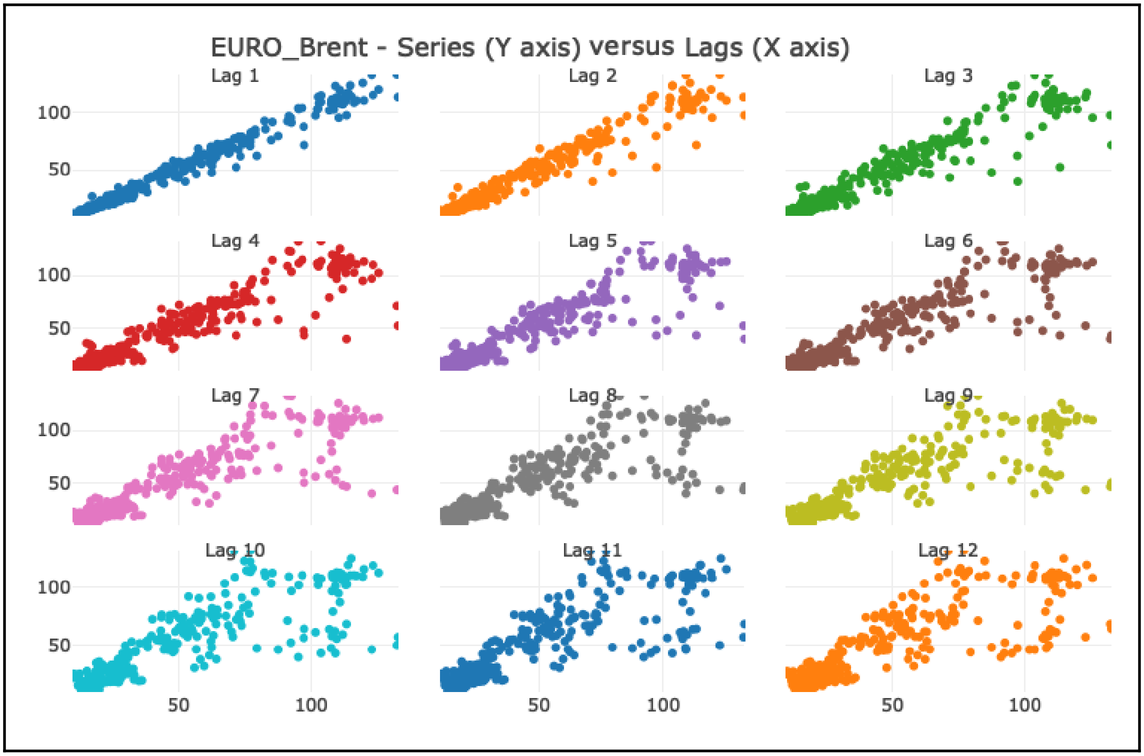

Series EURO_Brent
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
Lag

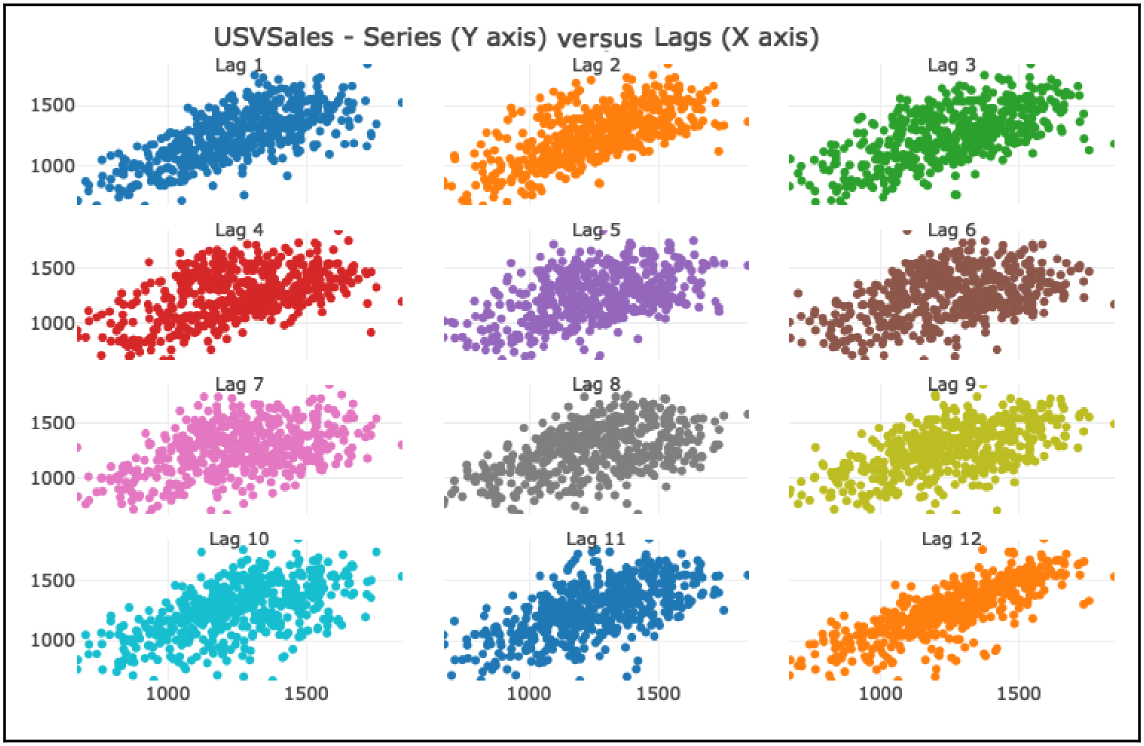

Series USVSales
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
1
2
3
4
5
Lag

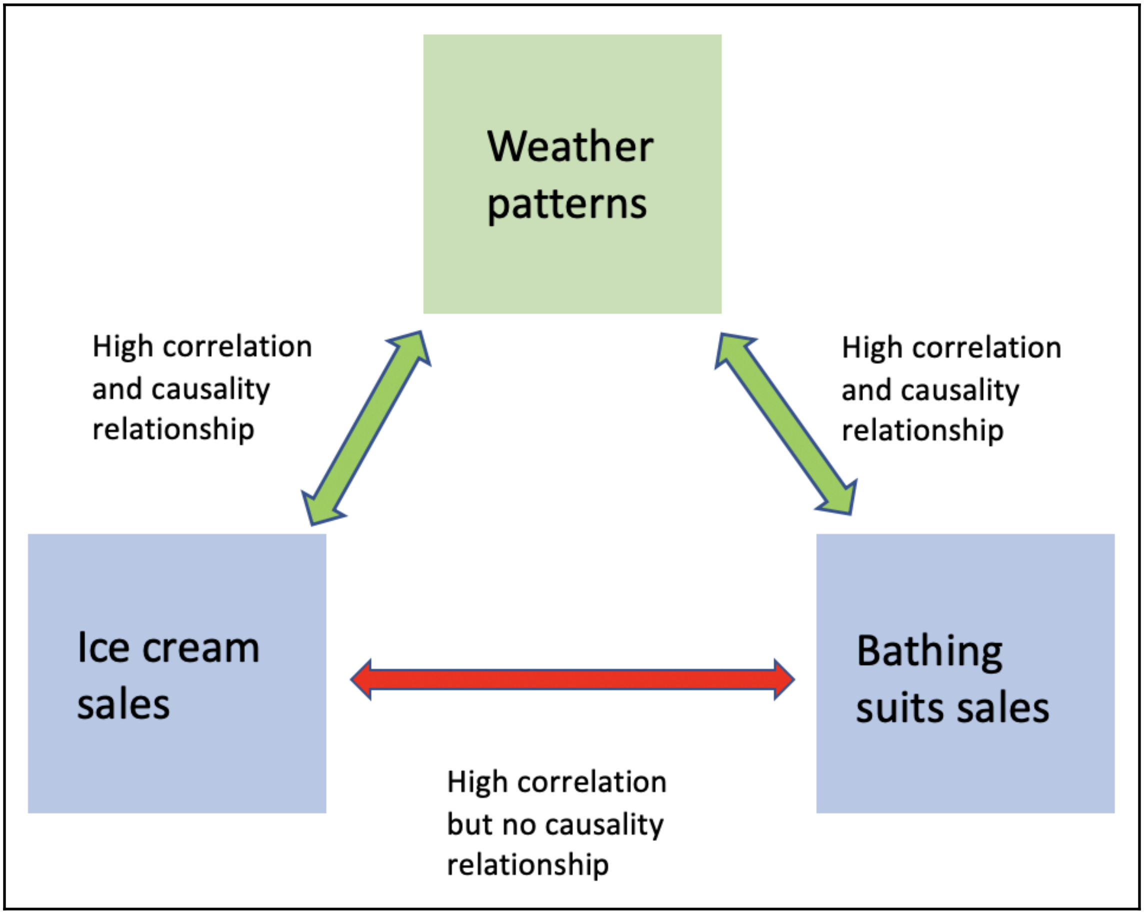

Series USgas
Partial ACF
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4
5
Lag

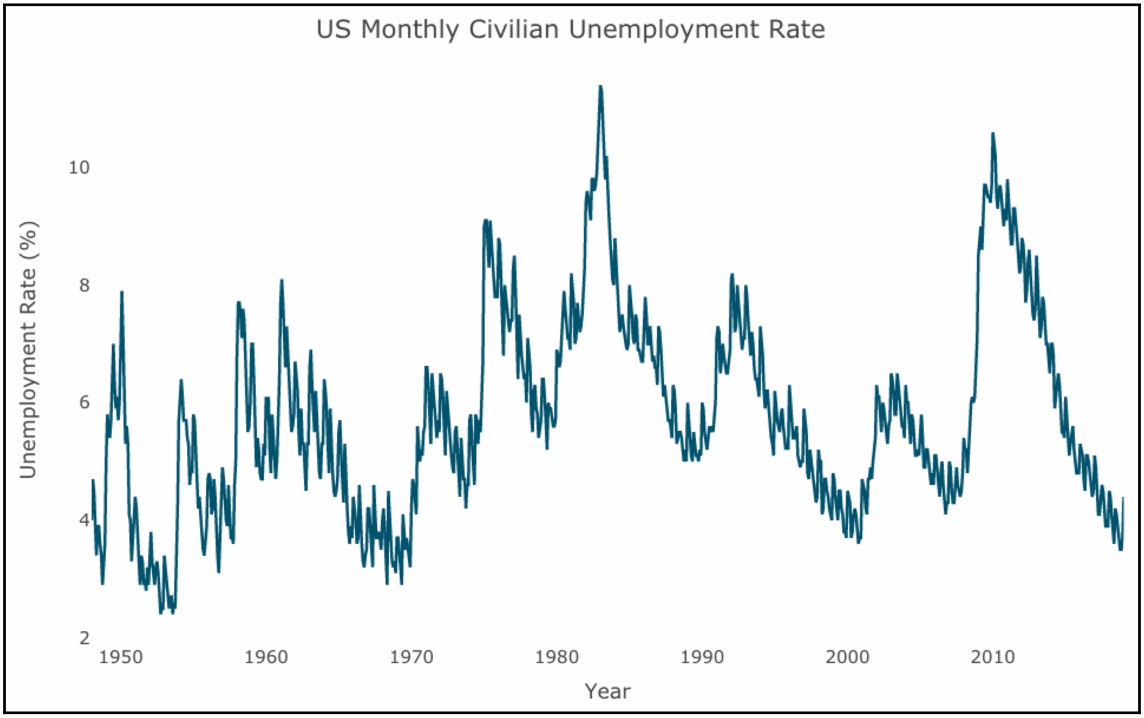

USgas - Series (Y axis) versus Lags (X axis)
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Lag 11
Lag 12
3000
2500
2000
1500
1500
2000
2500
3000

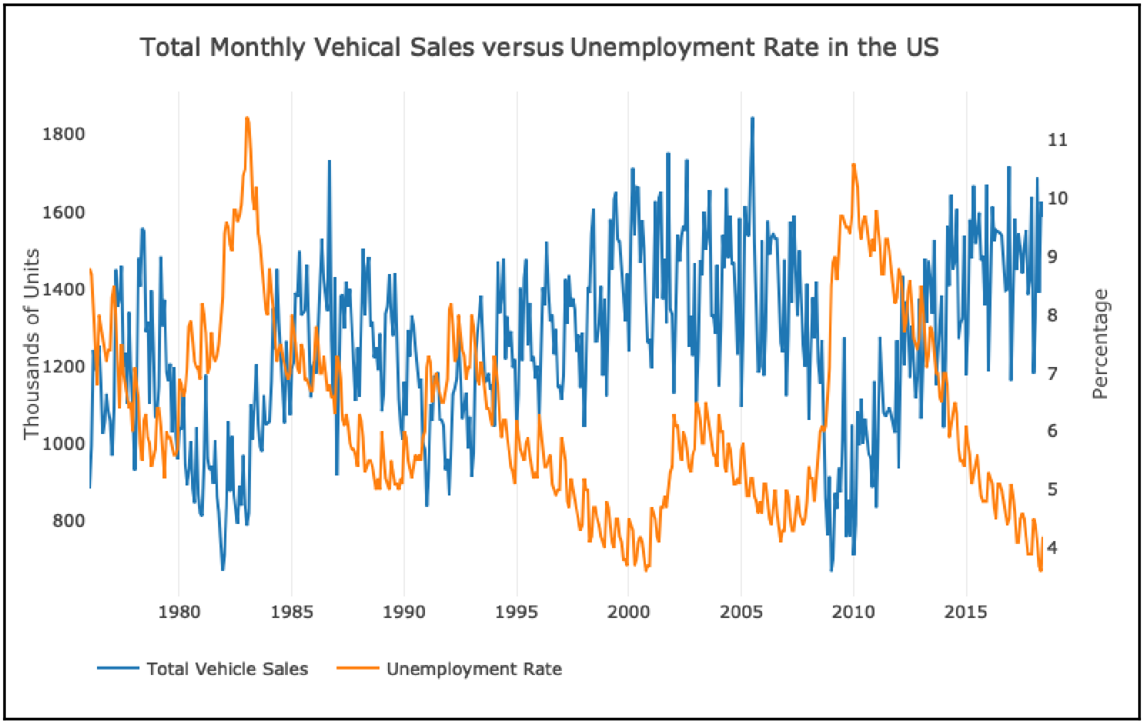

USgas - Series (Y axis) versus Lags (X axis)
Lag 12
Lag 24
Lag 36
Lag 48
3000
2500
2000
1500
1500
2000
2500
3000

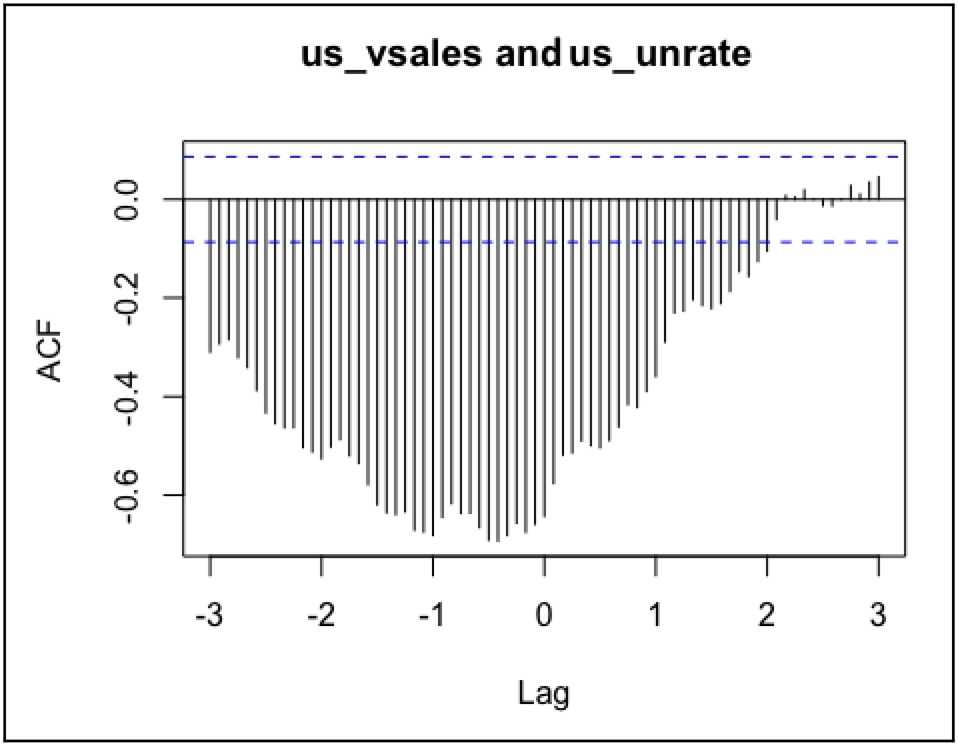

EURO_Brent - Series (Y axis) versus Lags (X axis)
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Lag 11
Lag 12
100
50
50
100

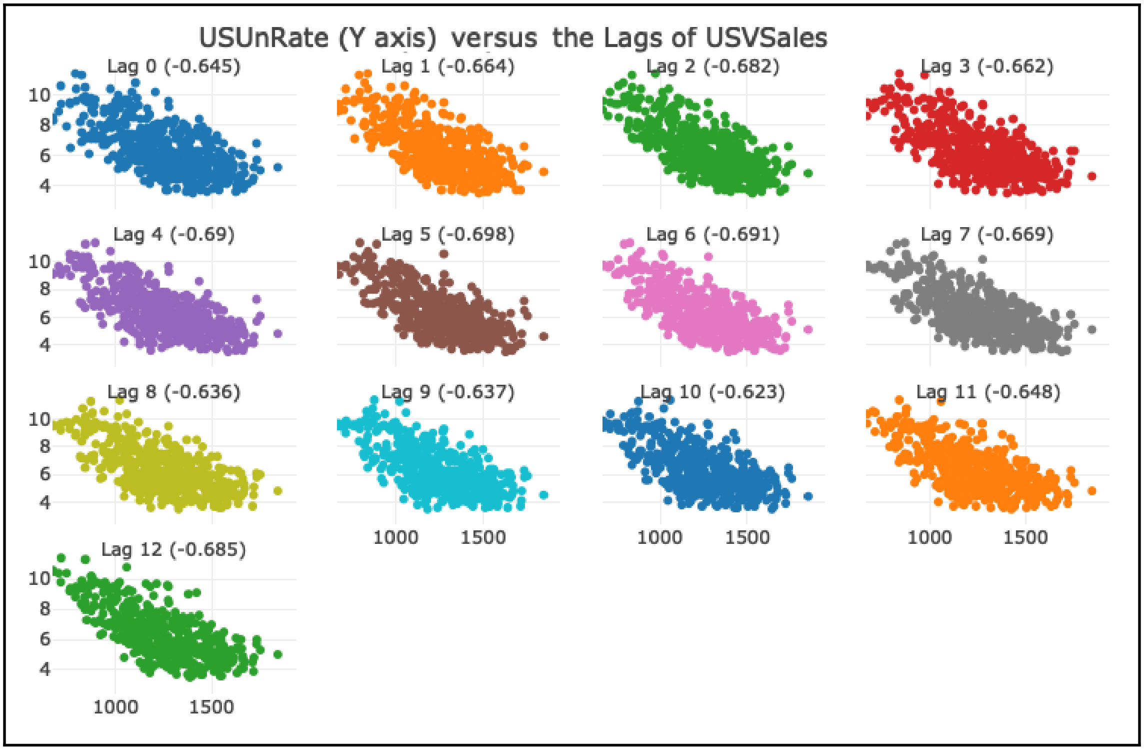

USVSales - Series (Y axis) versus Lags (X axis)
Lag 1
Lag 2
Lag 3
Lag 4
Lag 5
Lag 6
Lag 7
Lag 8
Lag 9
Lag 10
Lag 11
Lag 12
1500
1000
1000
1500

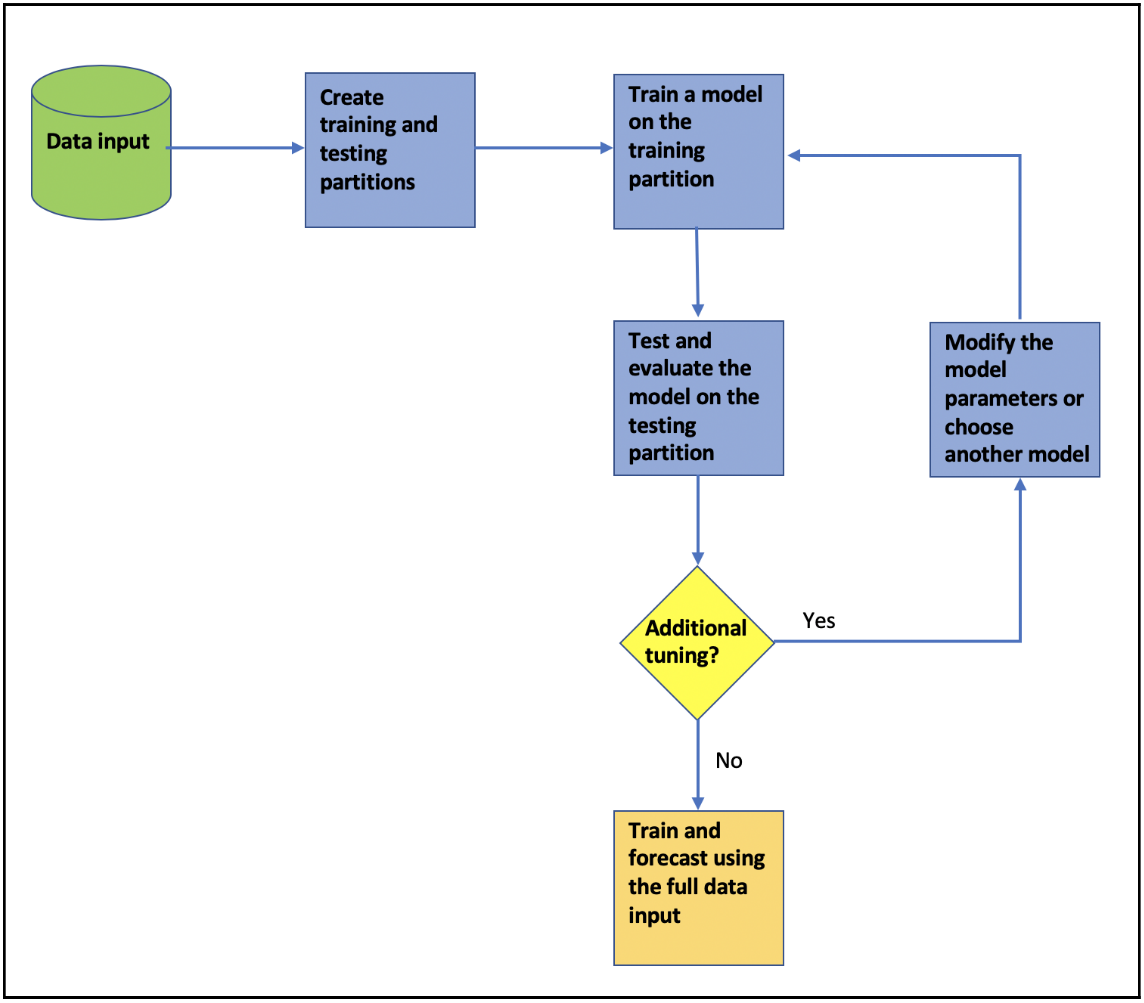

Weather patterns
High correlation and causality relationship
High correlation and causality relationship
Ice cream sales
Bathing suits sales
High correlation but no causality relationship

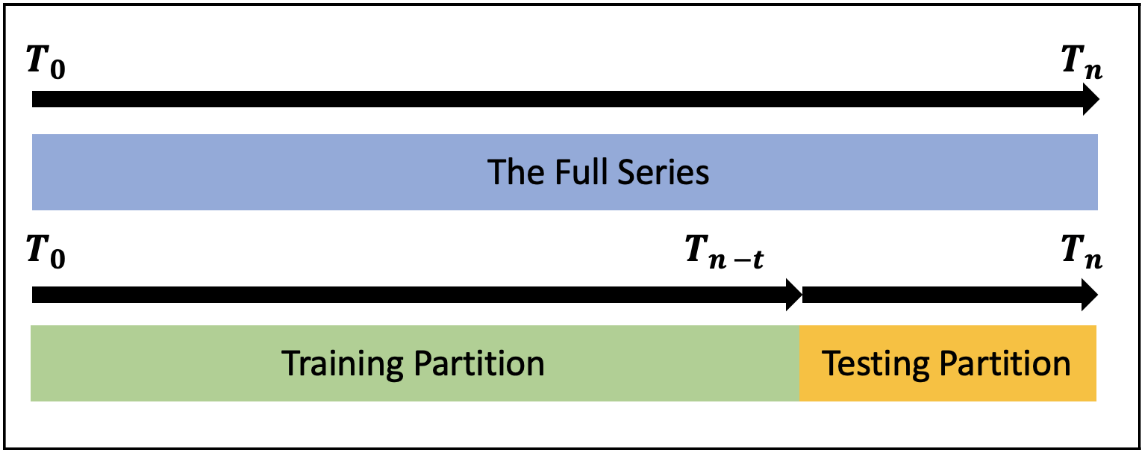

US Monthly Civilian Unemployment Rate
Unemployment Rate (%)
2
4
6
8
10
1950
1960
1970
1980
1990
2000
2010
Year

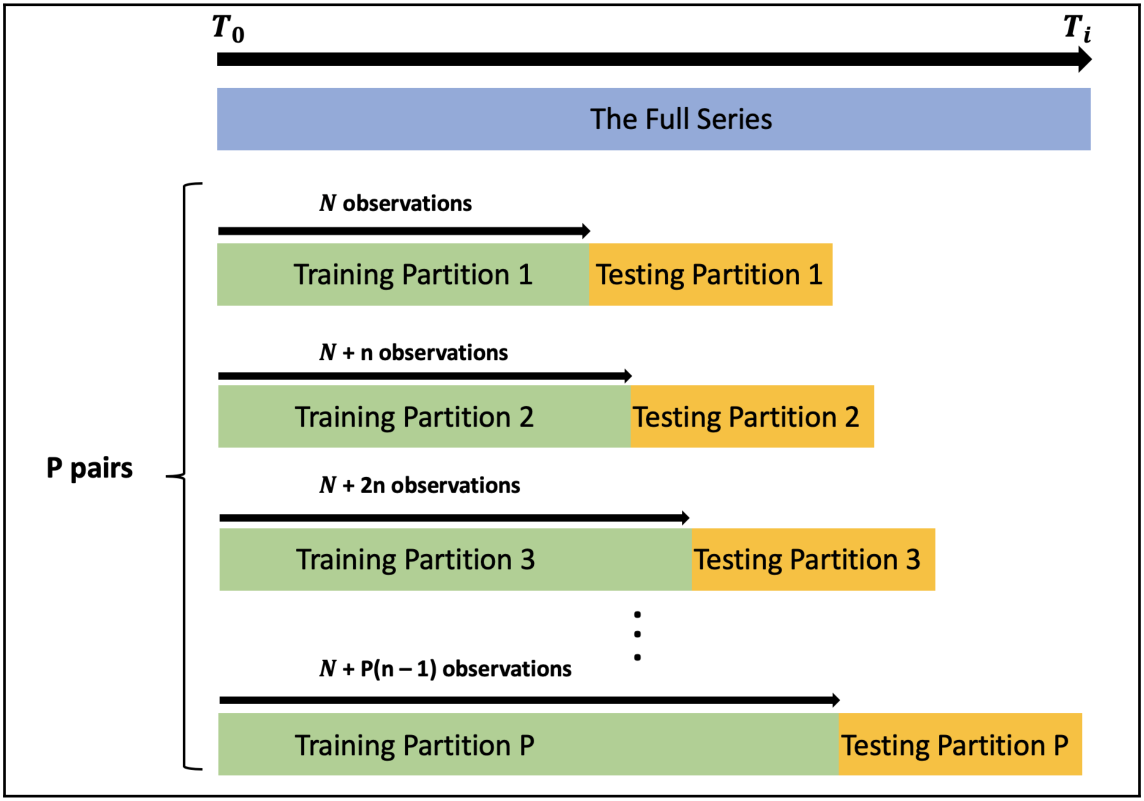

Total Monthly Vehical Sales versus Unemployment Rate in the US
Thousands of Units
1800
1600
1400
1200
1000
800
Percentage
11
10
9
8
7
6
5
4
1980
1985
1990
1995
2000
2005
2010
2015
Total Vehicle Sales
Unemployment Rate

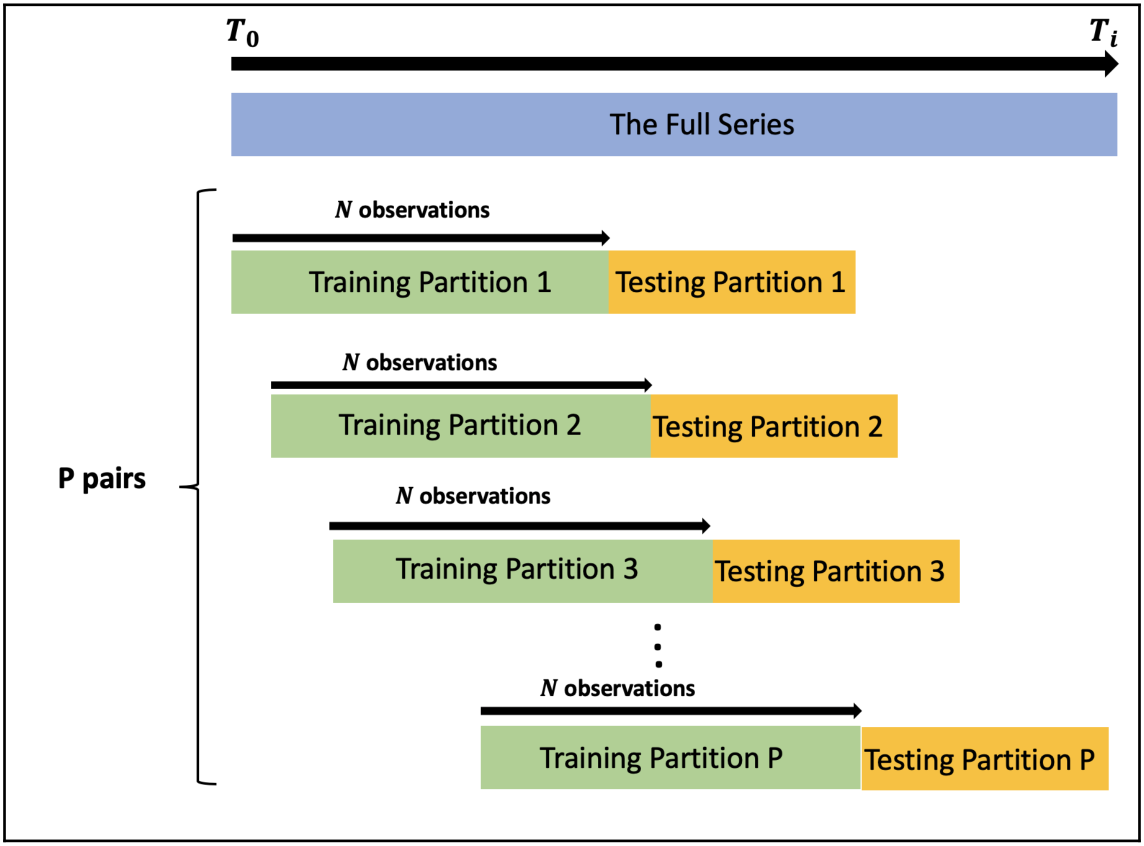

us_vsales and us_unrate
ACF
0.0
-0.2
-0.4
-0.6
-3
-2
-1
0
1
2
3
Lag

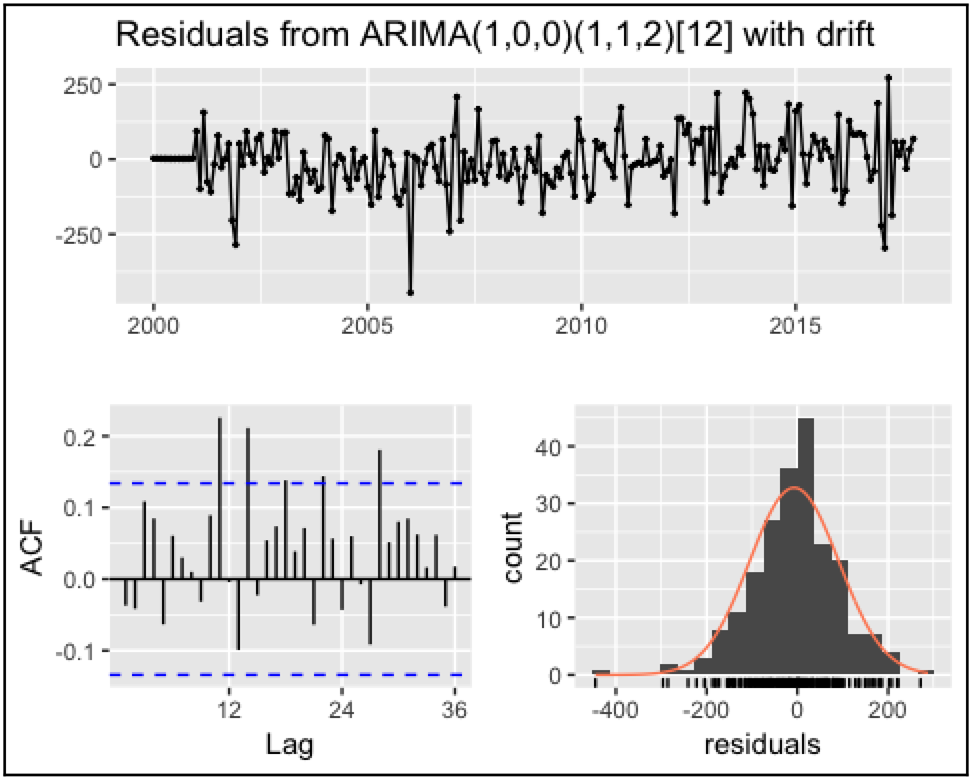

USUnRate (Y axis) versus the Lags of USVSales
Lag 0 (-0.645)
Lag 1 (-0.664)
Lag 2 (-0.682)
Lag 3 (-0.662)
Lag 4 (-0.69)
Lag 5 (-0.698)
Lag 6 (-0.691)
Lag 7 (-0.669)
Lag 8 (-0.636)
Lag 9 (-0.637)
Lag 10 (-0.623)
Lag 11 (-0.648)
Lag 12 (-0.685)
10
8
6
4
1000
1500

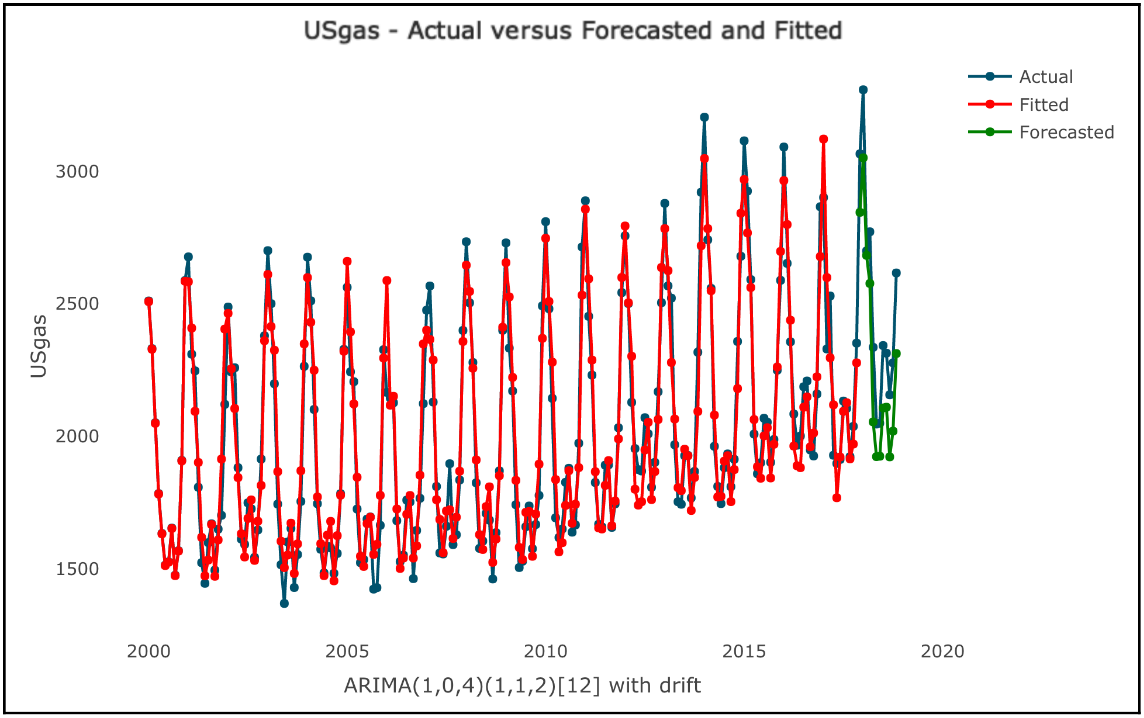

Data input
Create training and testing partitions
Train a model on the training partition
Test and evaluate the model on the testing partition
Modify the model parameters or choose another model
Additional tuning?
Yes
No
Train and forecast using the full data input

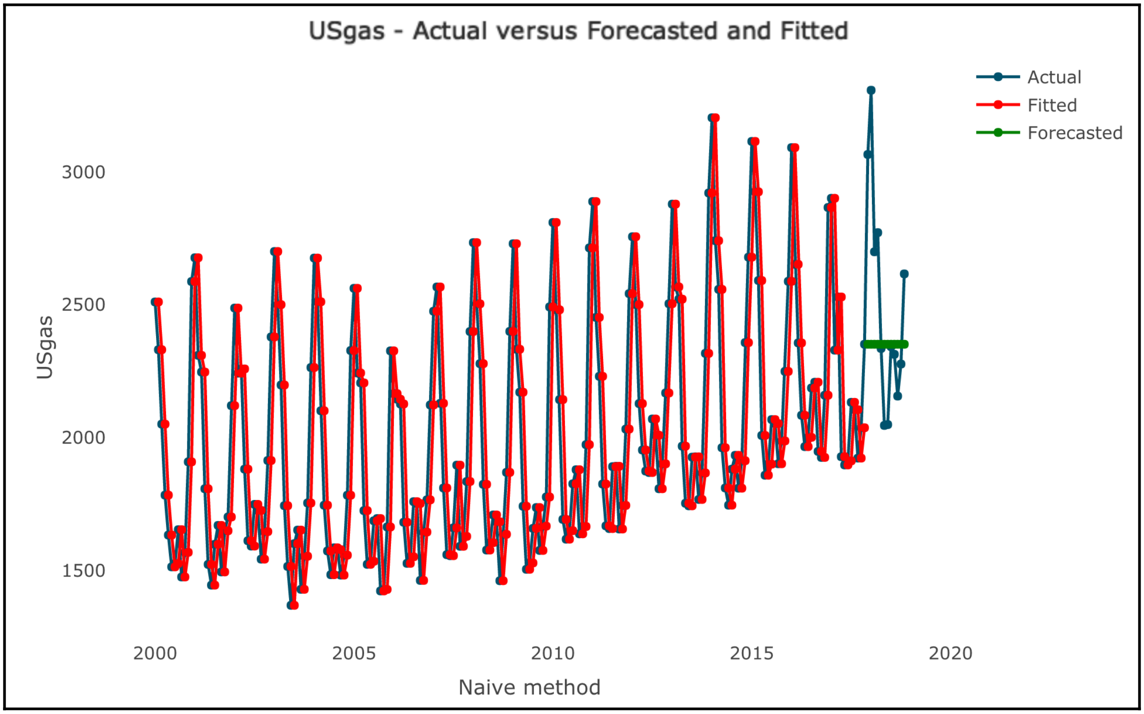

T_0
T_n
The Full Series
T_0
T_{n-t}
T_n
Training Partition
Testing Partition

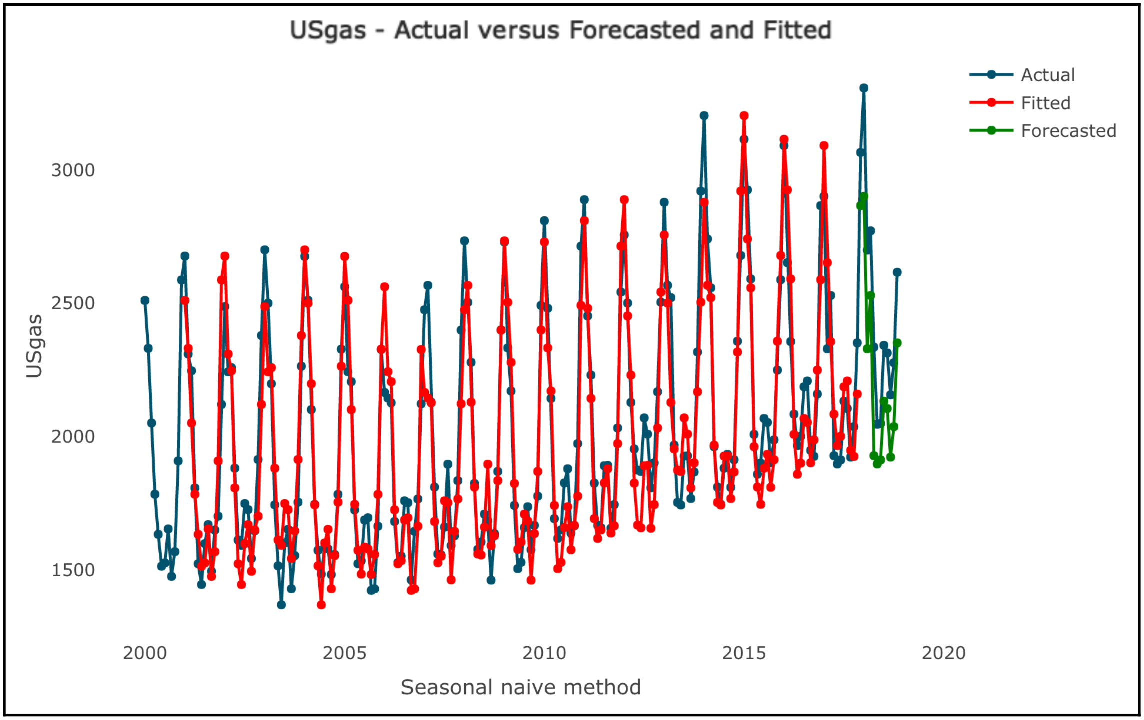

T_0
T_i
The Full Series
N observations
Training Partition 1
Testing Partition 1
N + n observations
Training Partition 2
Testing Partition 2
N + 2n observations
Training Partition 3
Testing Partition 3
N + P(n – 1) observations
Training Partition P
Testing Partition P
P pairs

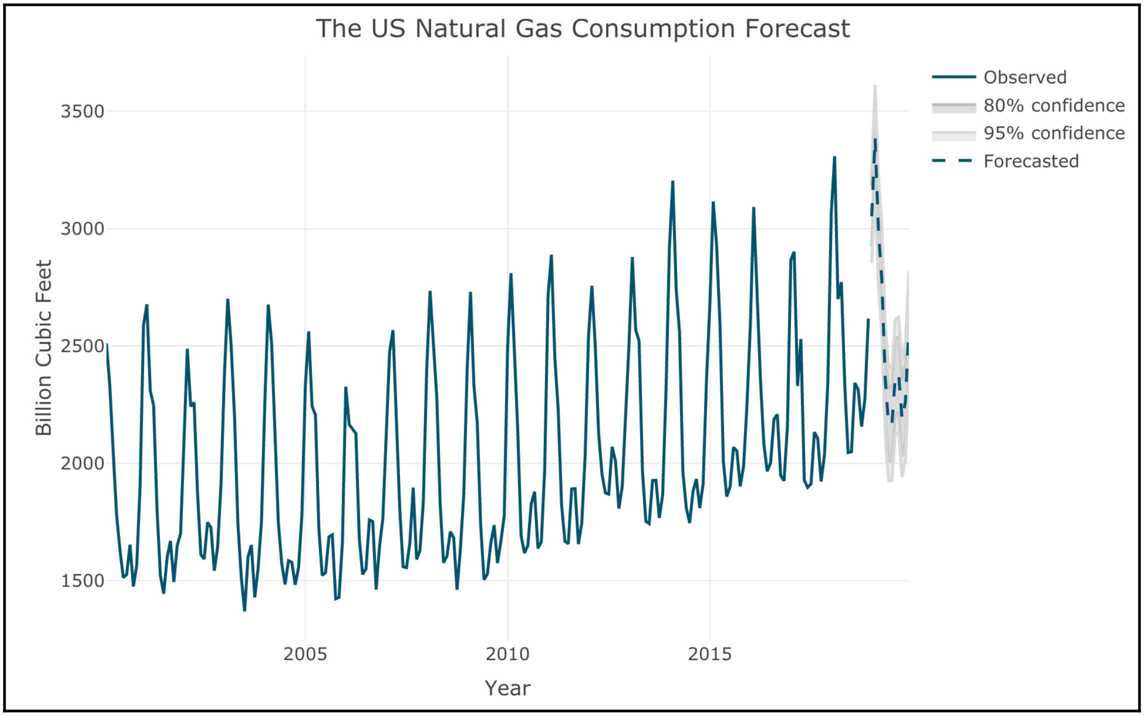

T_0
T_i
The Full Series
P pairs
N observations
Training Partition 1
Testing Partition 1
N observations
Training Partition 2
Testing Partition 2
N observations
Training Partition 3
Testing Partition 3
N observations
Training Partition P
Testing Partition P

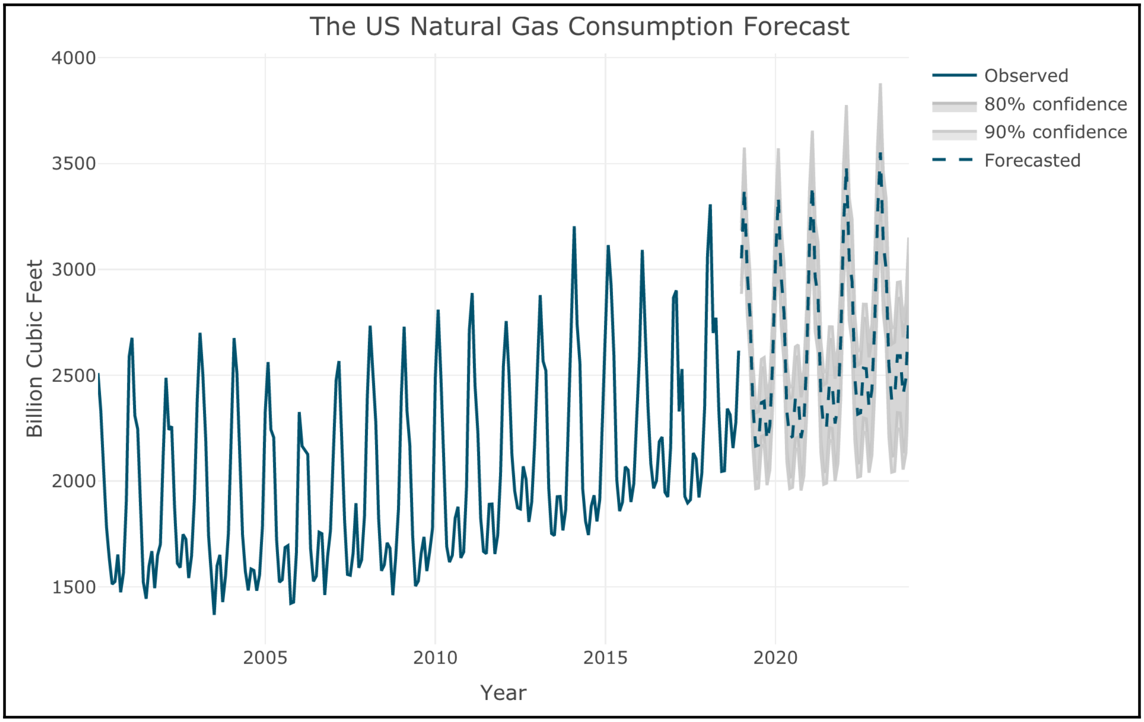

Residuals from ARIMA(1,0,0)(1,1,2)[12] with drift
250
0
-250
2000
2005
2010
2015
0.2
0.1
0.0
-0.1
ACF
12
24
36
Lag
40
30
20
10
0
count
-400
-200
0
200
residuals

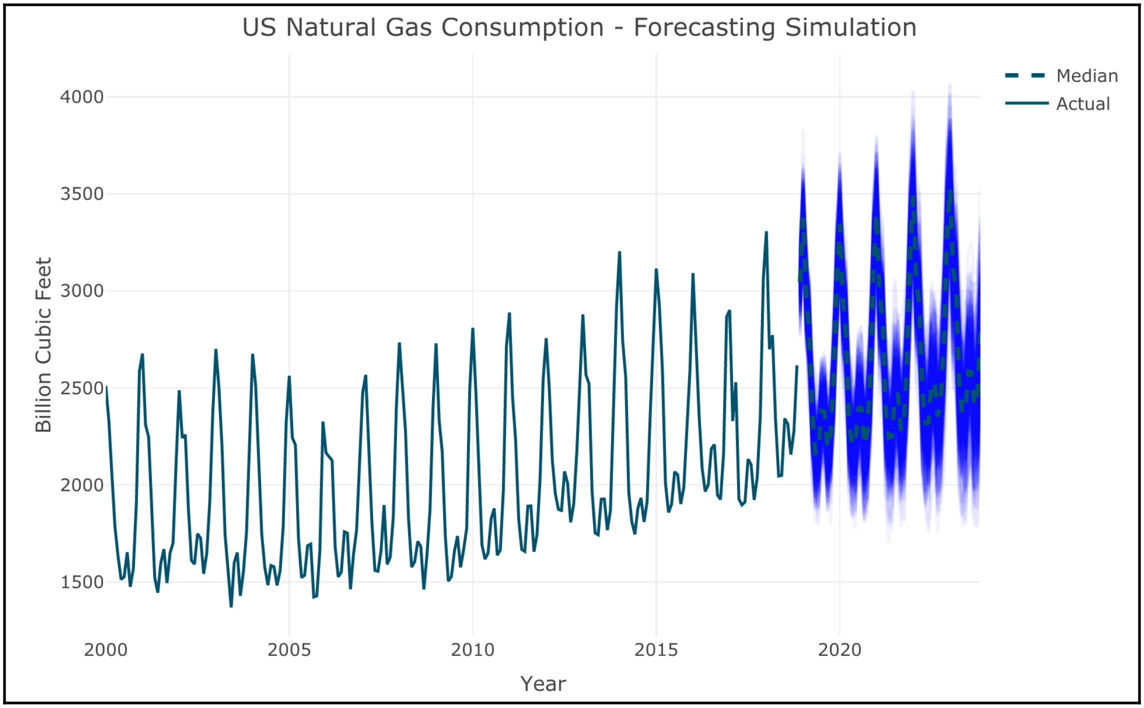

USgas - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
3000
2500
2000
1500
USgas
2000
2005
2010
2015
2020
ARIMA(1,0,4)(1,1,2)[12] with drift


USgas - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
3000
2500
2000
1500
USgas
2000
2005
2010
2015
2020
Naive method

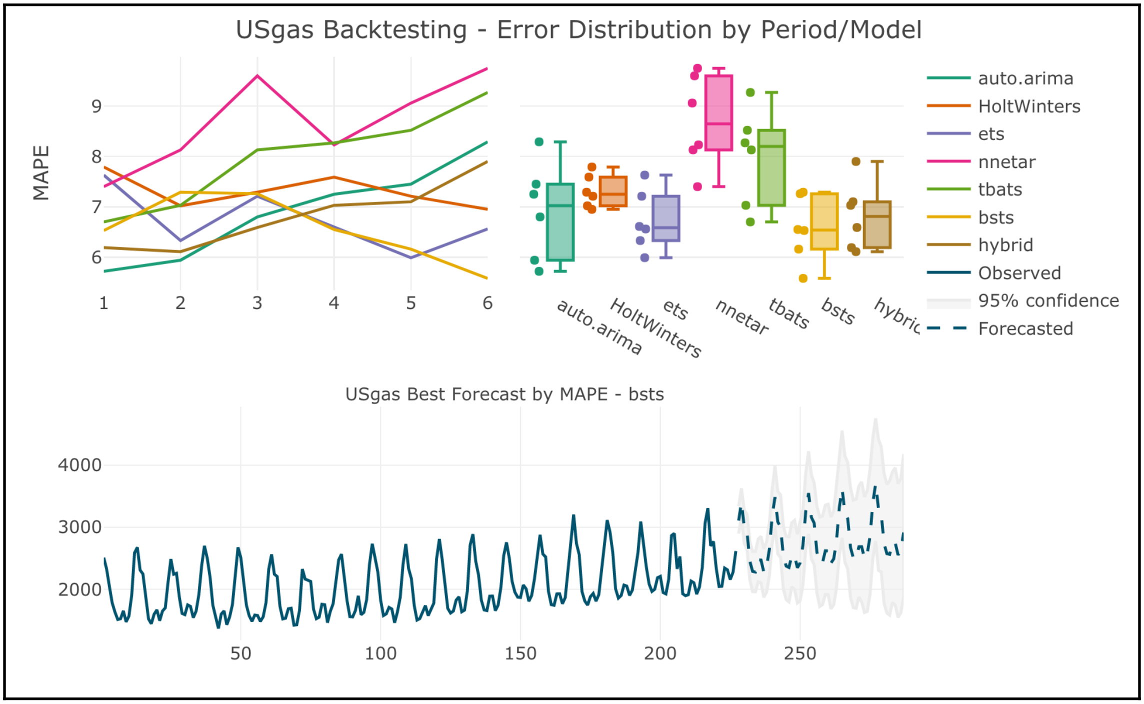

USgas - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
3000
2500
2000
1500
USgas
2000
2005
2010
2015
2020
Seasonal naive method


The US Natural Gas Consumption Forecast
Observed
80% confidence
95% confidence
Forecasted
Billion Cubic Feet
3500
3000
2500
2000
1500
2005
2010
2015
Year


The US Natural Gas Consumption Forecast
4000
3500
3000
2500
2000
1500
Billion Cubic Feet
2005
2010
2015
2020
Year
Observed
80% confidence
90% confidence
Forecasted

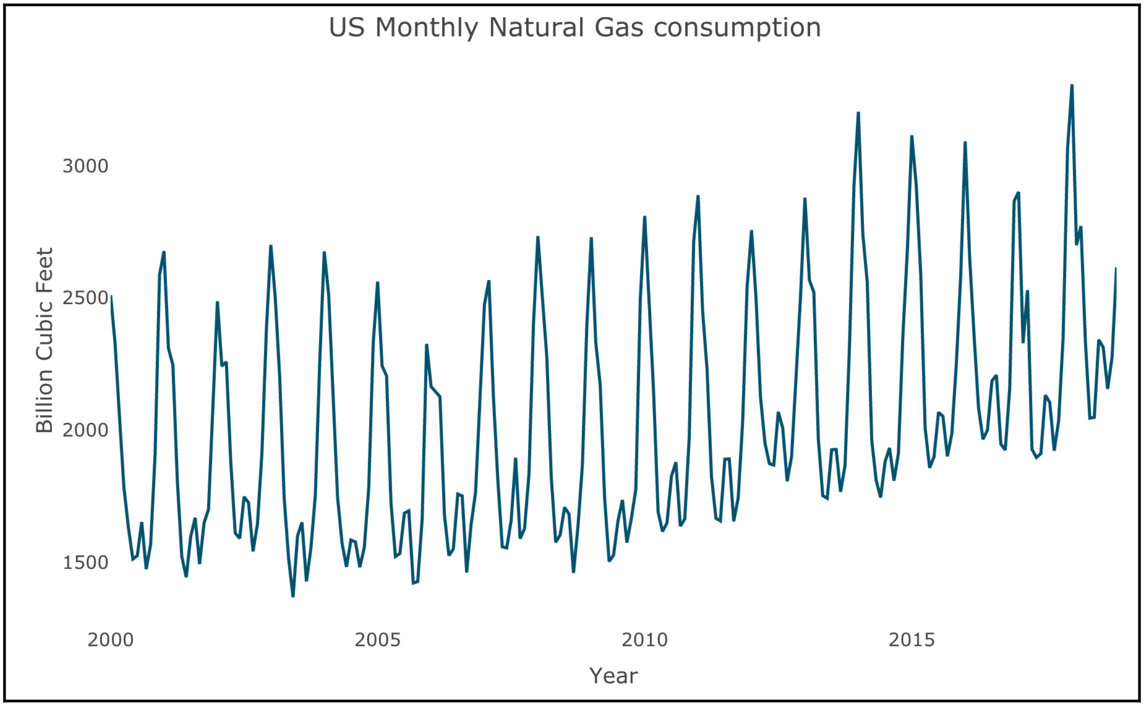


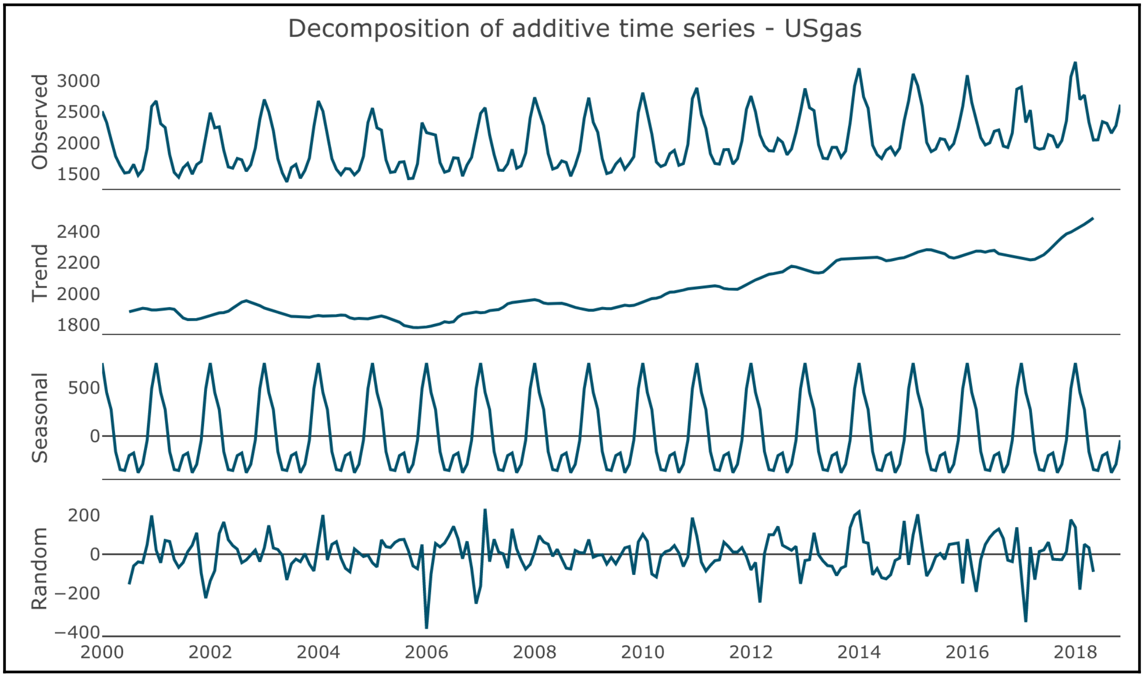
US Natural Gas Consumption - Forecasting Simulation
Median
Actual
Billion Cubic Feet
4000
3500
3000
2500
2000
1500
2000
2005
2010
2015
2020
Year

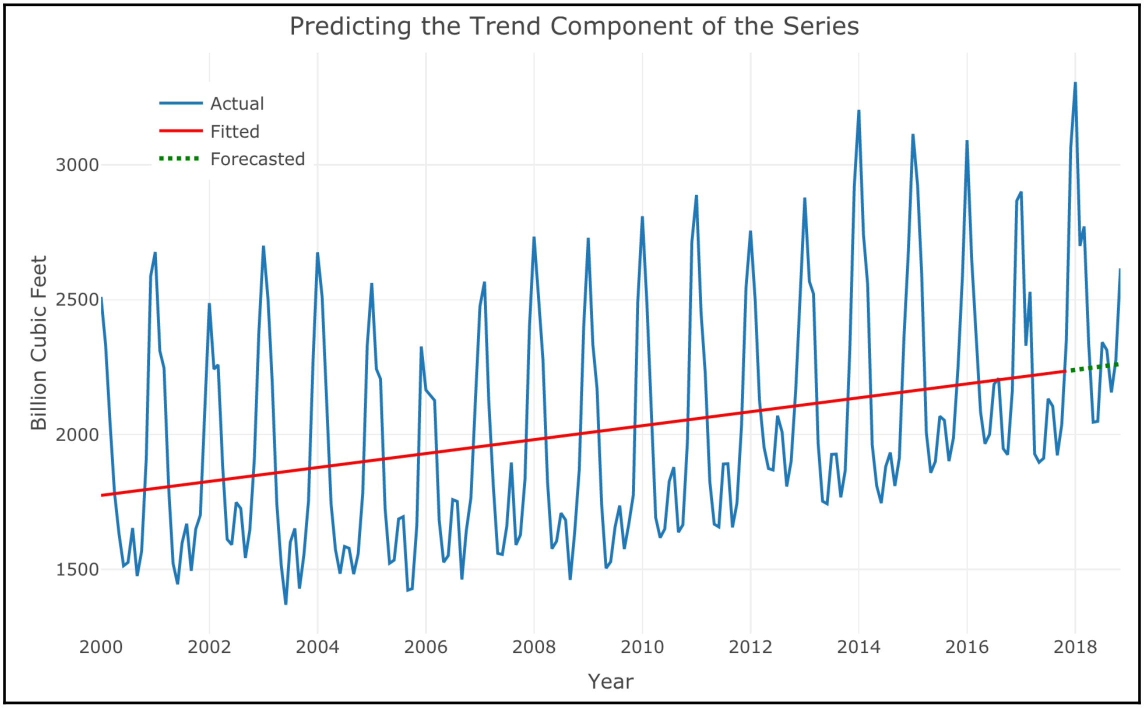


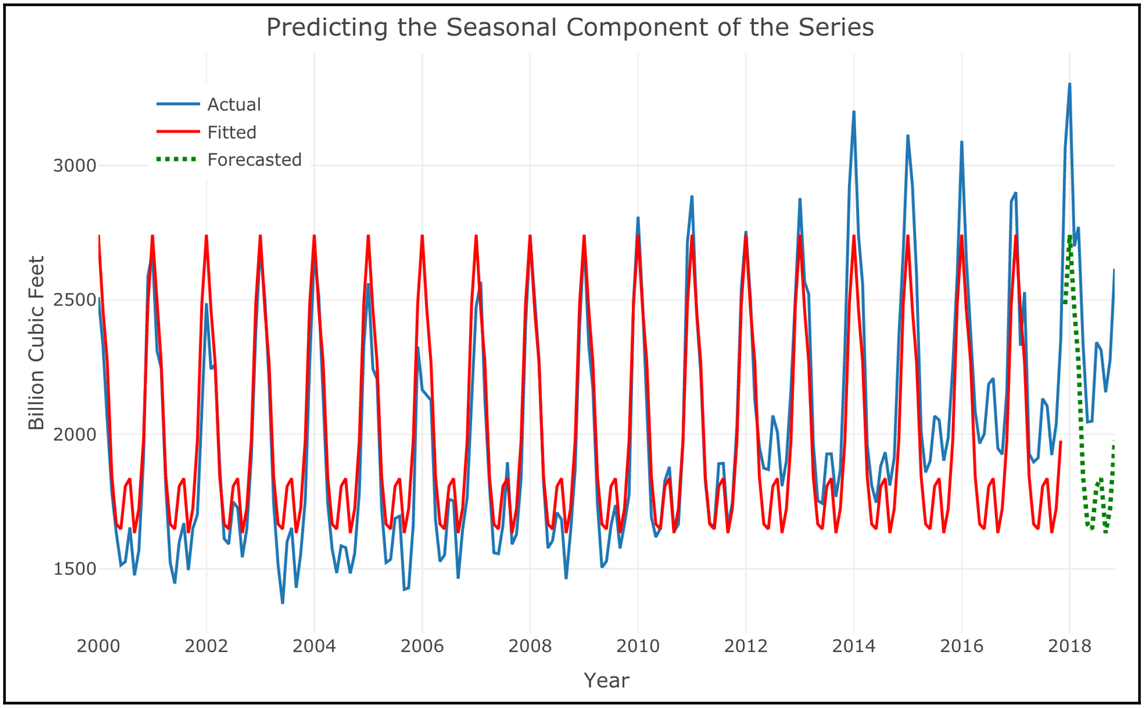
USgas Backtesting - Error Distribution by Period/Model
MAPE
9
8
7
6
1
2
3
4
5
6
auto.arima
HoltWinters
ets
nnetar
tbats
bsts
hybrid
auto.arima
HoltWinters
ets
nnetar
tbats
bsts
hybrid
Observed
95% confidence
Forecasted
USgas Best Forecast by MAPE - bsts
4000
3000
2000
50
100
150
200
250

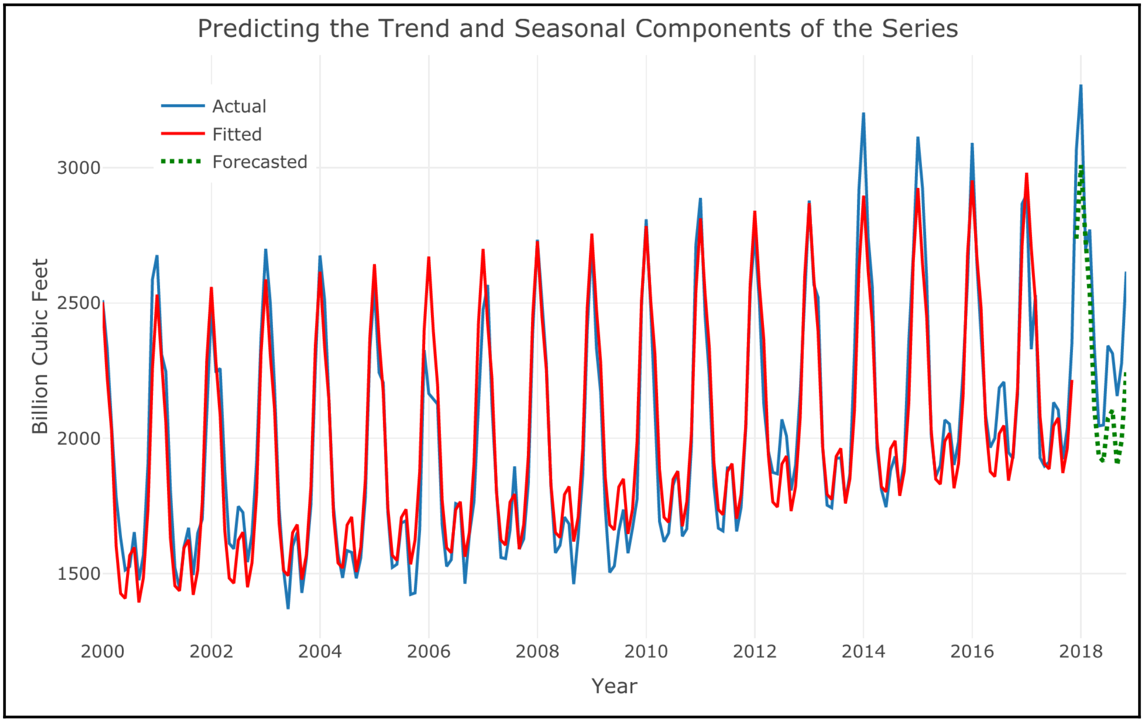
Chapter 9: Forecasting with Linear Regression

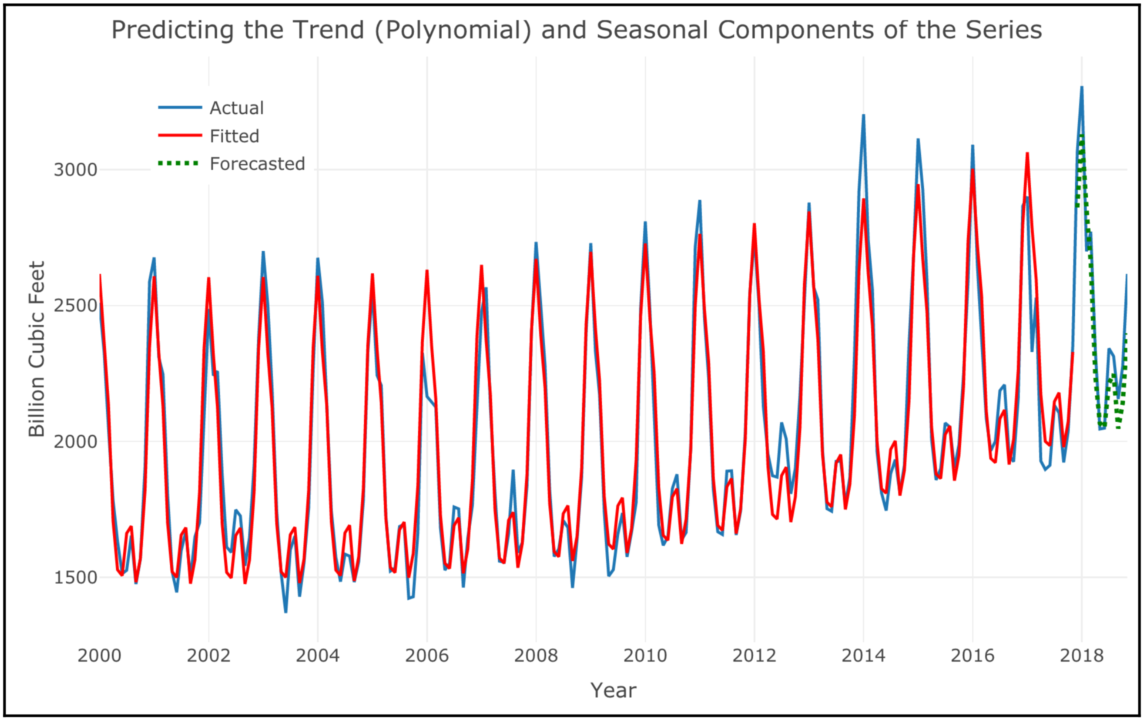


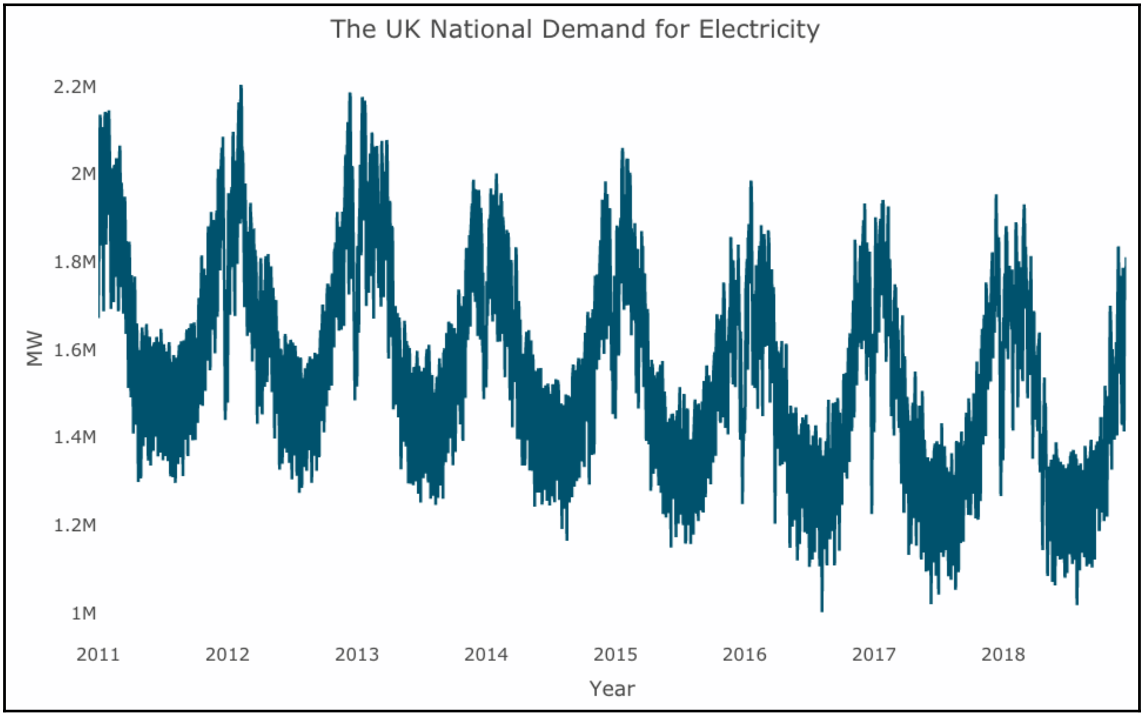

Decomposition of additive time series - USgas
Observed
3000
2500
2000
1500
Trend
2400
2200
2000
1800
Seasonal
500
0
Random
200
0
−200
−400
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018

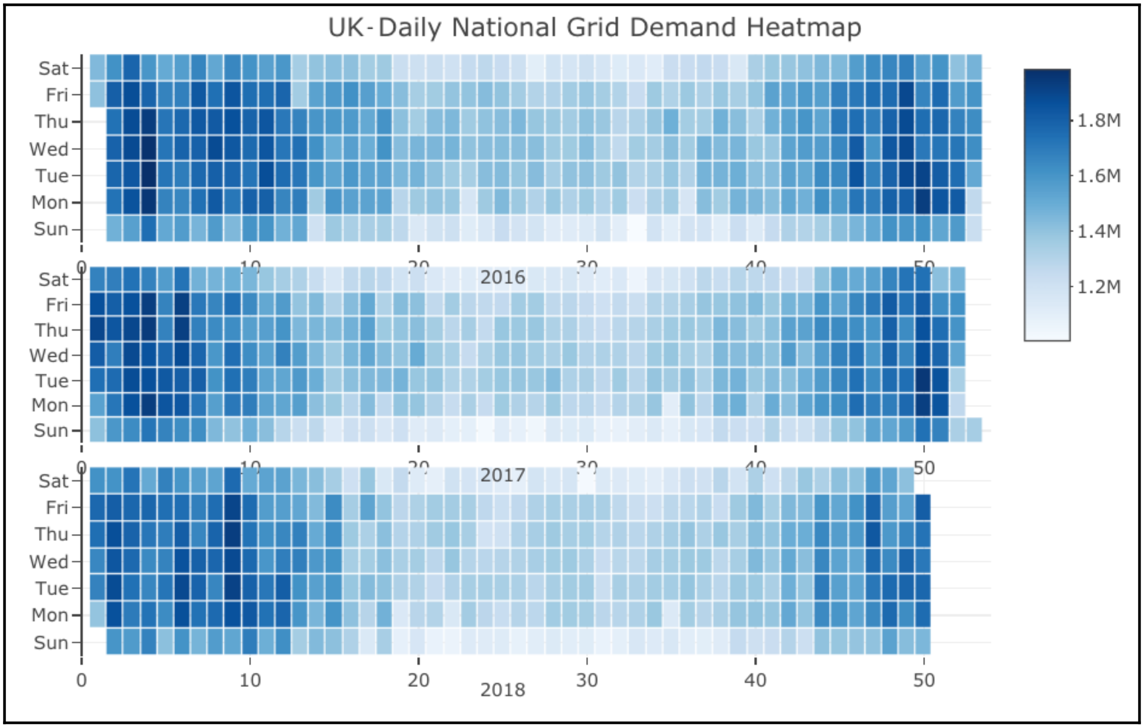

Predicting the Trend Component of the Series
Actual
Fitted
Forecasted
Billion Cubic Feet
3000
2500
2000
1500
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year

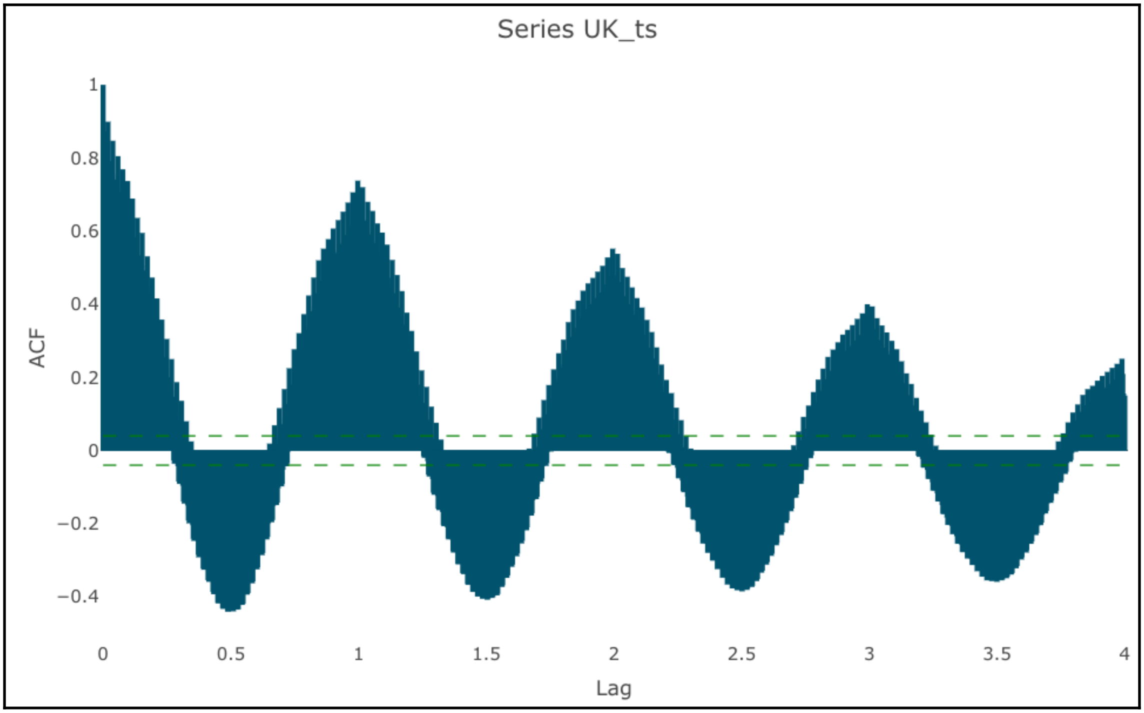

Predicting the Seasonal Component of the Series
Actual
Fitted
Forecasted
Billion Cubic Feet
3000
2500
2000
1500
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year

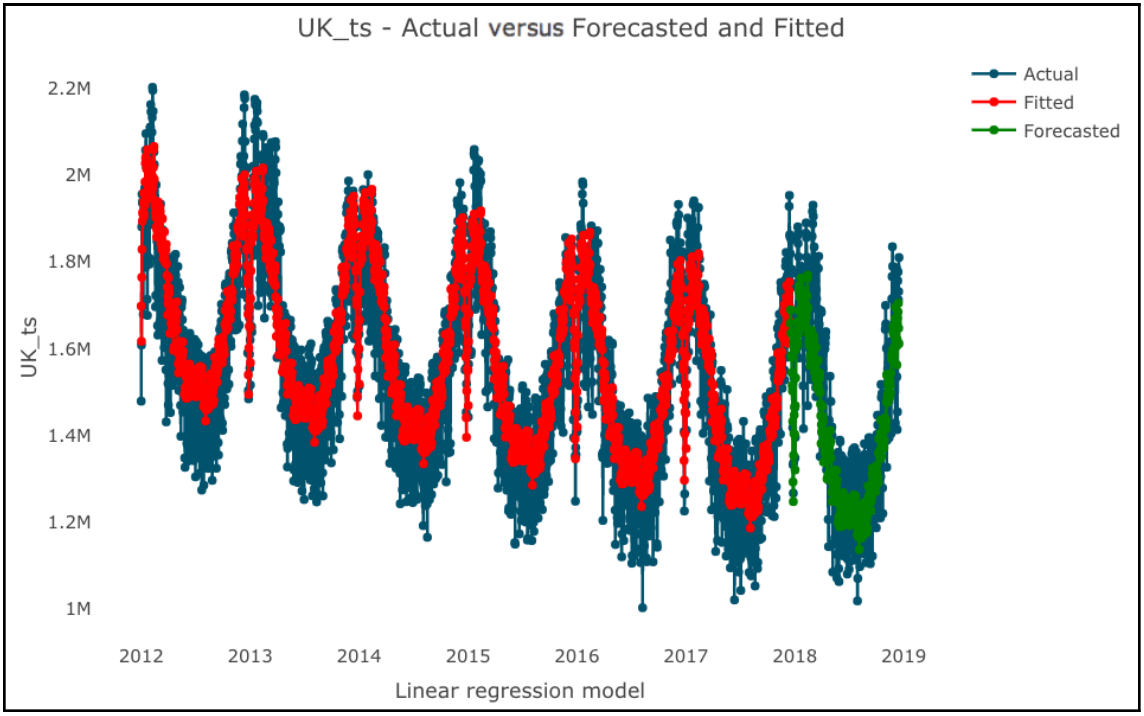

Predicting the Trend and Seasonal Components of the Series
Actual
Fitted
Forecasted
3000
2500
2000
1500
Billion Cubic Feet
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year

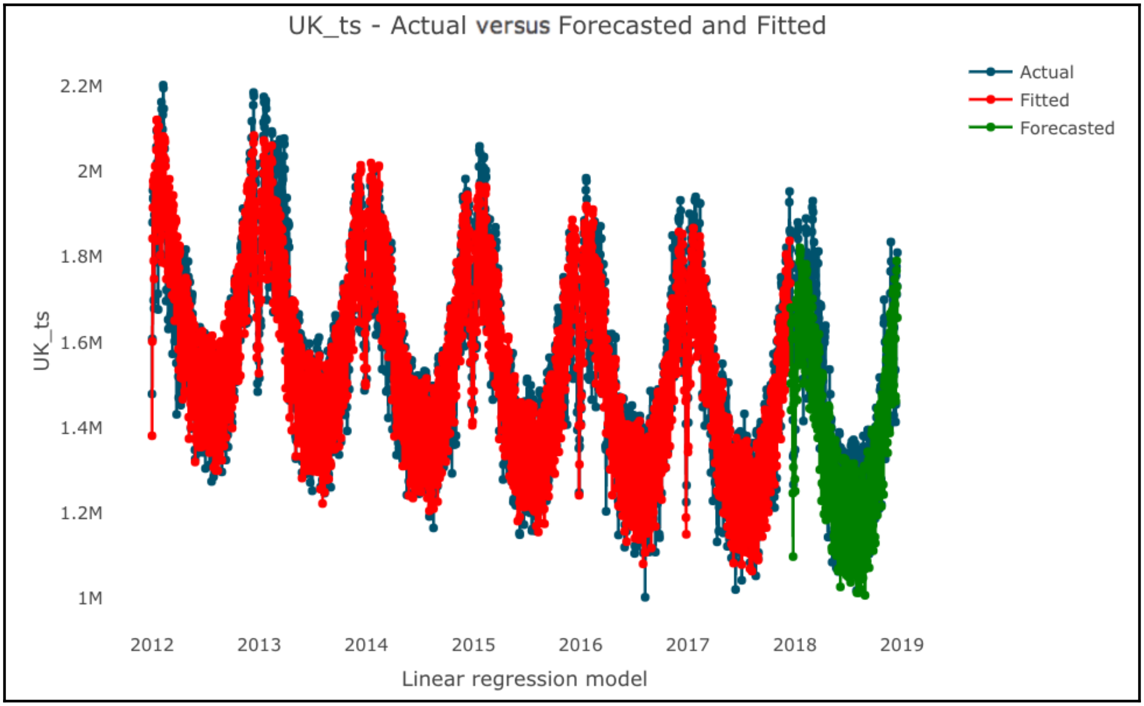

Predicting the Trend (Polynomial) and Seasonal Components of the Series
Actual
Fitted
Forecasted
Billion Cubic Feet
3000
2500
2000
1500
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
Year

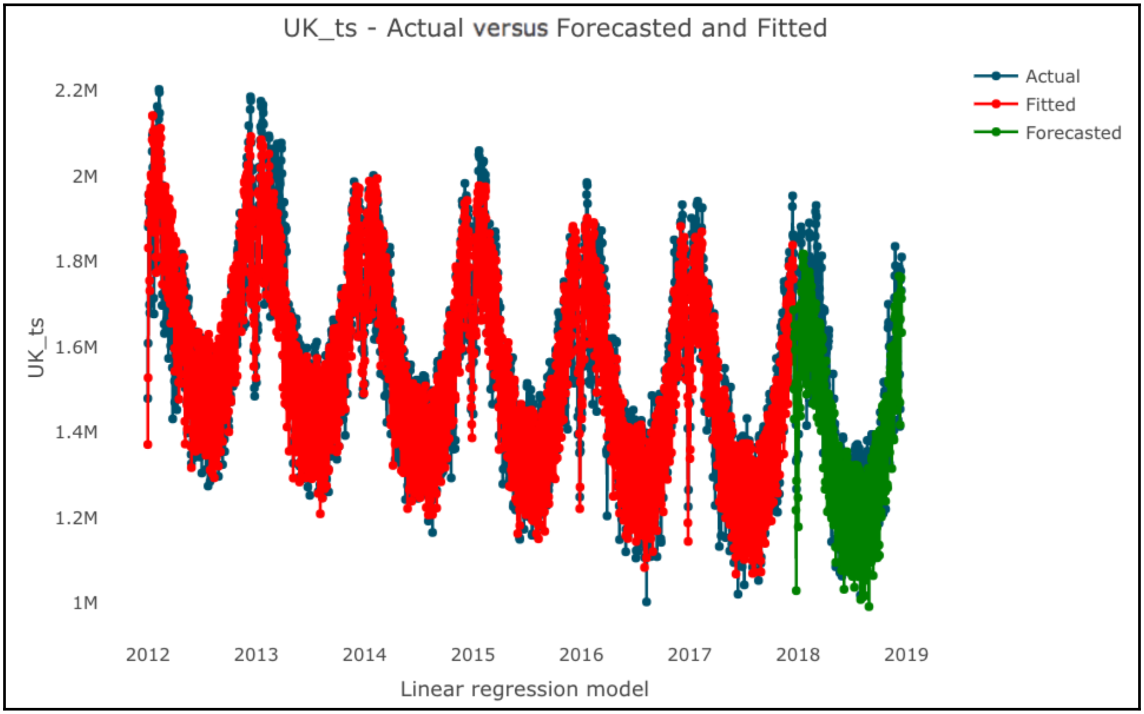

The UK National Demand for Electricity
2.2M
2M
1.8M
1.6M
1.4M
1.2M
1M
MW
2011
2012
2013
2014
2015
2016
2017
2018
Year

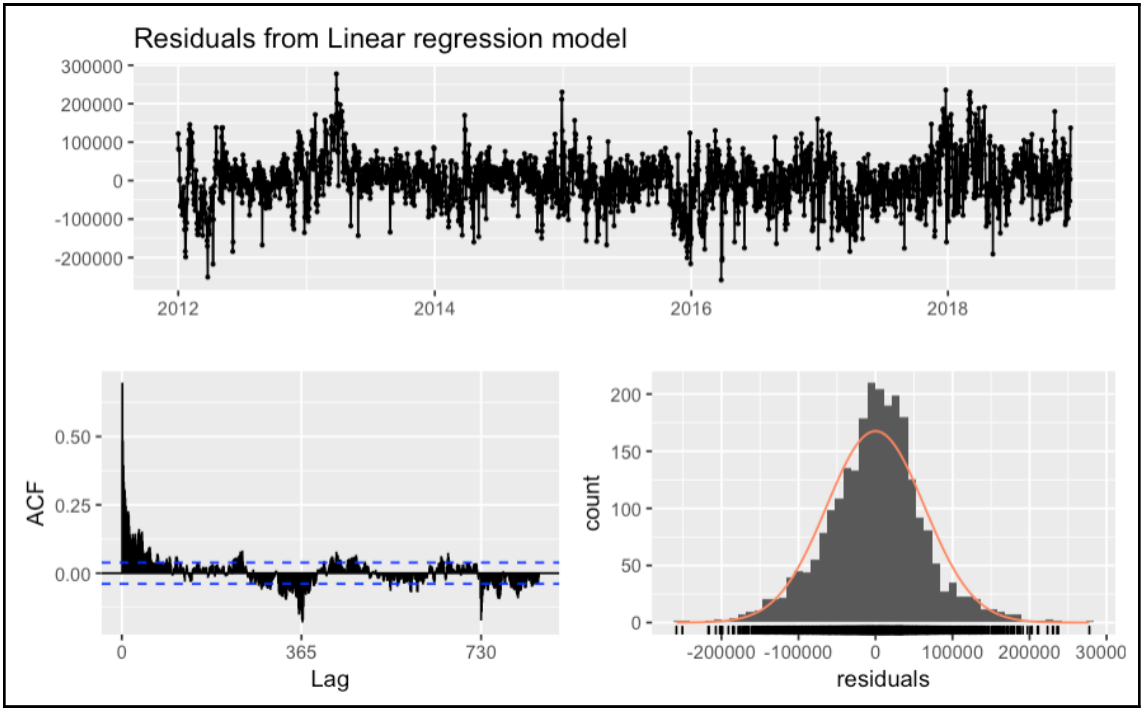

UK-Daily National Grid Demand Heatmap
Sat
Fri
Thu
Wed
Tue
Mon
Sun
0
10
20
30
40
50
2016
2017
2018
1.8M
1.6M
1.4M
1.2M

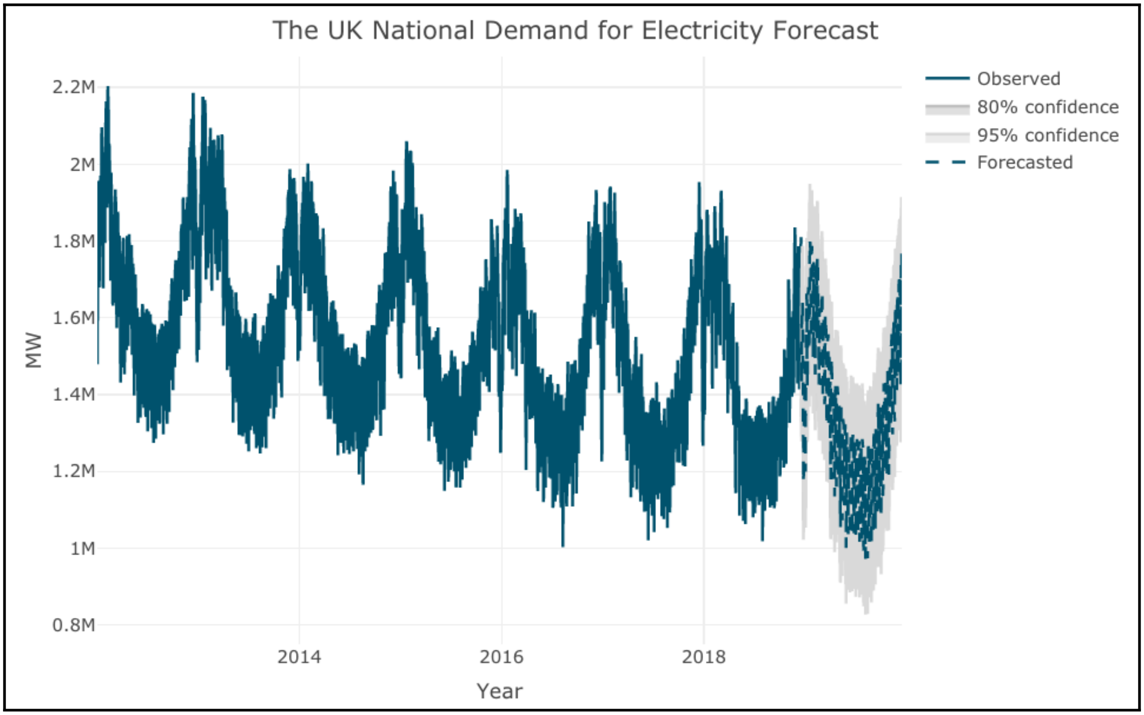

Series UK_ts
ACF
Lag
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
0
0.5
1
1.5
2
2.5
3
3.5
4


UK_ts - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
2.2M
2M
1.8M
1.6M
1.4M
1.2M
1M
UK_ts
2012
2013
2014
2015
2016
2017
2018
2019
Linear regression model

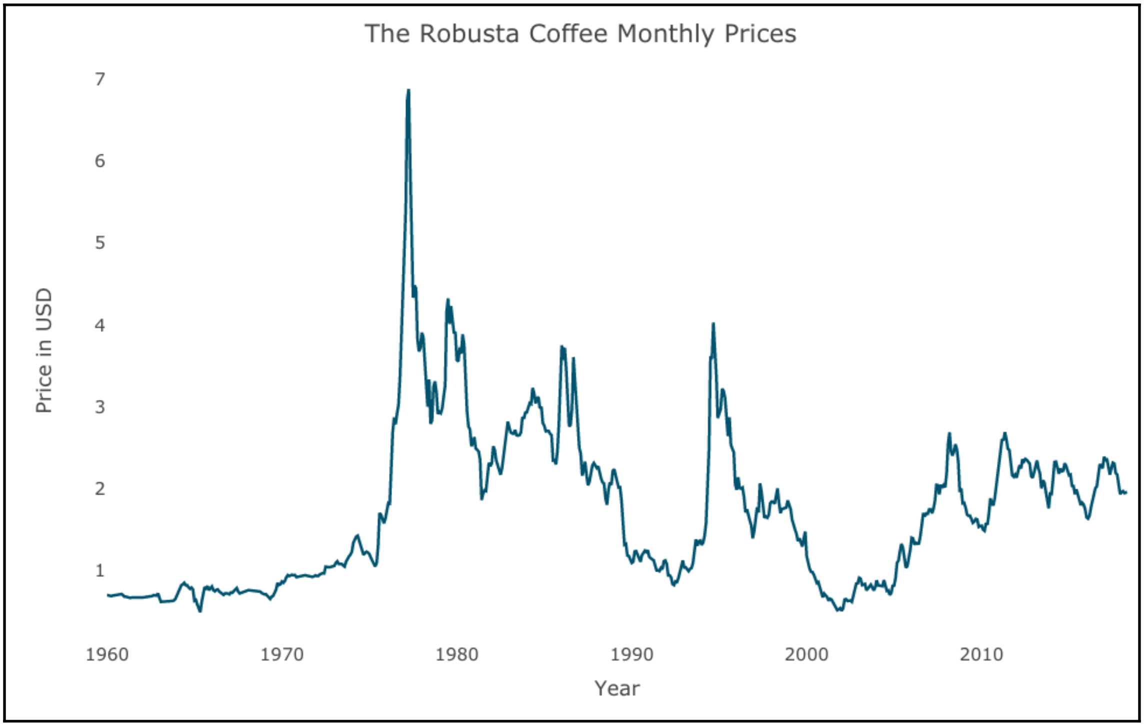

UK_ts - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
2.2M
2M
1.8M
1.6M
1.4M
1.2M
1M
UK_ts
2012
2013
2014
2015
2016
2017
2018
2019
Linear regression model

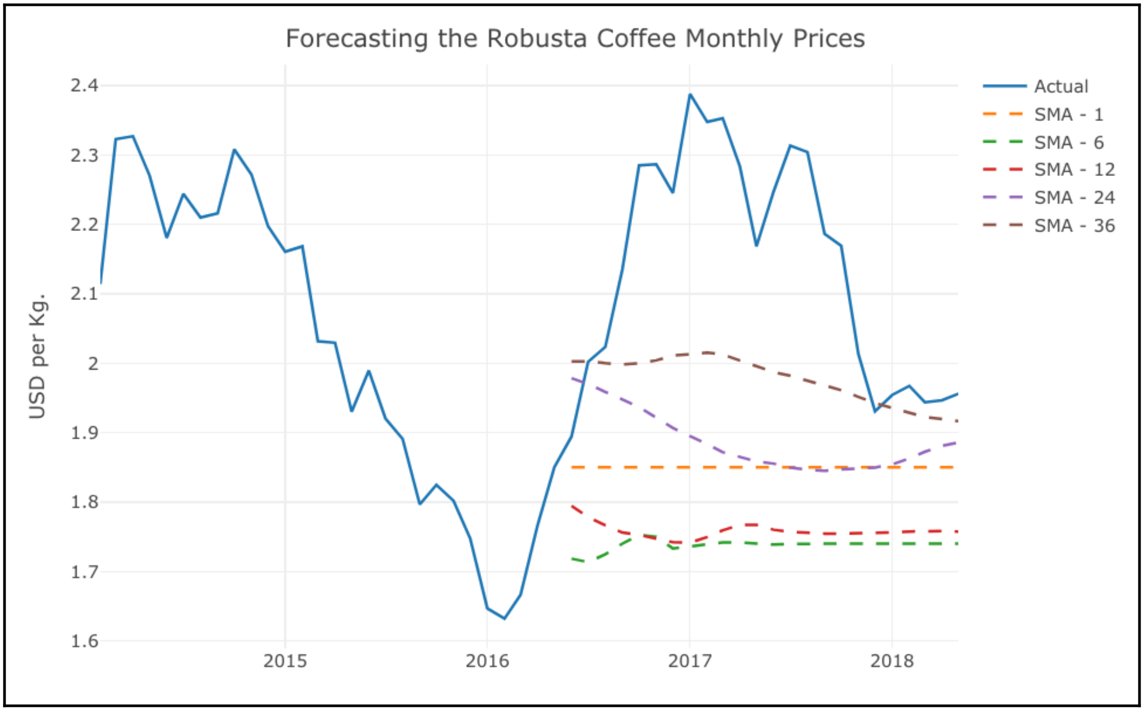

UK_ts - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
UK_ts
2.2M
2M
1.8M
1.6M
1.4M
1.2M
1M
2012
2013
2014
2015
2016
2017
2018
2019
Linear regression model

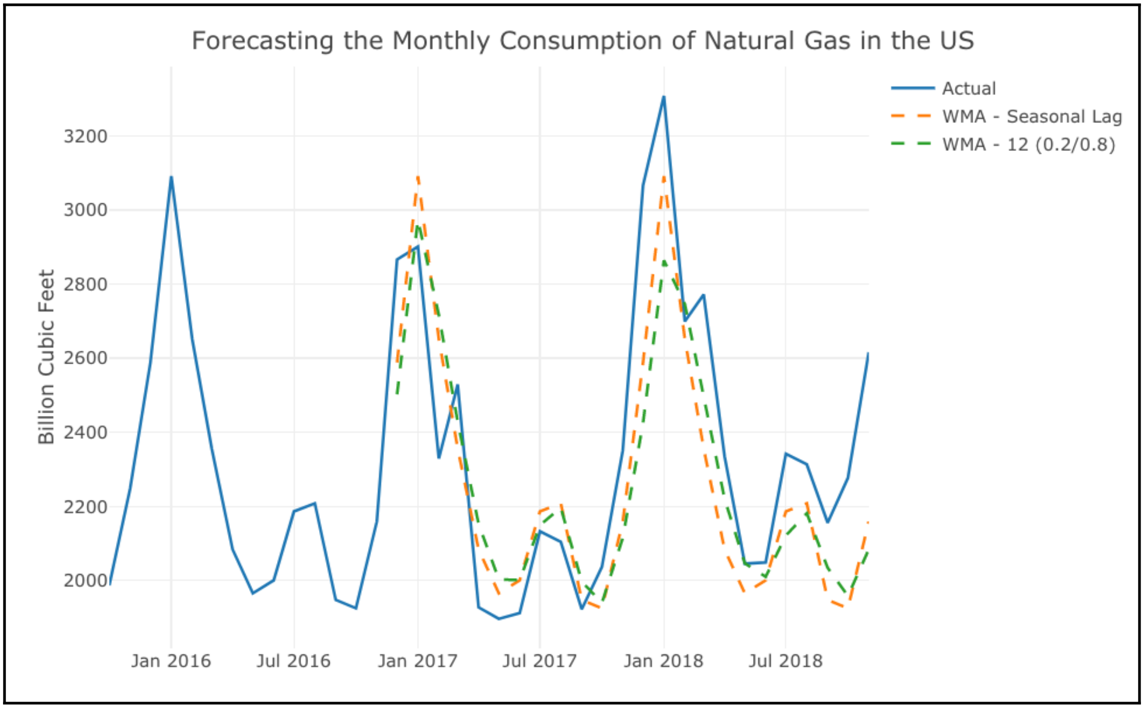

Residuals from Linear regression model
300000
200000
100000
0
-100000
-200000
2012
2014
2016
2018
ACF
0.50
0.25
0.00
0
365
730
Lag
count
200
150
100
50
0
-200000
-100000
0
100000
200000
30000
residuals


The UK National Demand for Electricity Forecast
Observed
80% confidence
95% confidence
Forecasted
2.2M
2M
1.8M
1.6M
1.4M
1.2M
1M
0.8M
MW
2014
2016
2018
Year

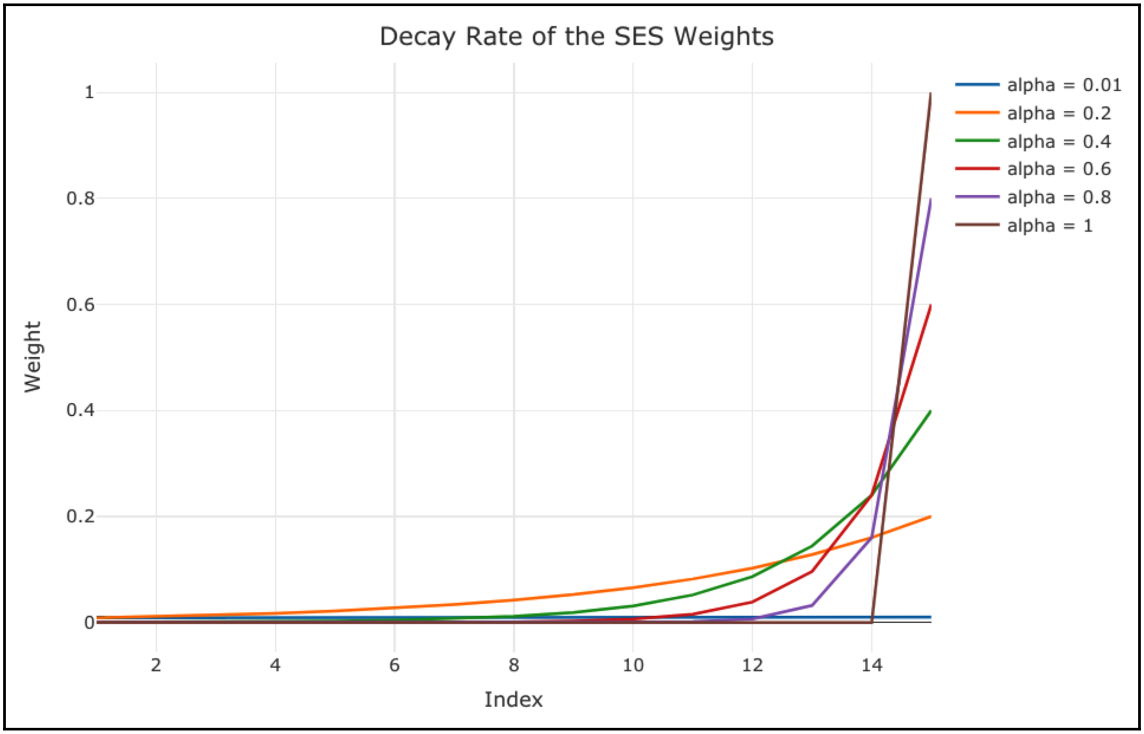
Chapter 10: Forecasting with Exponential Smoothing Models

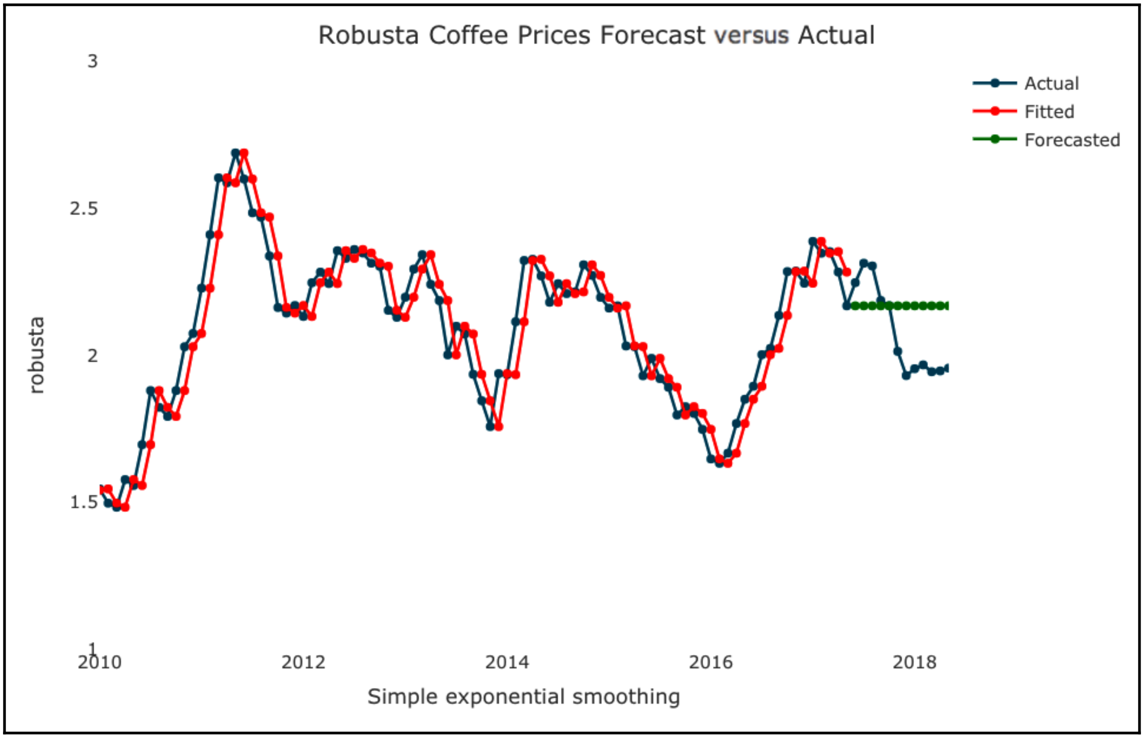


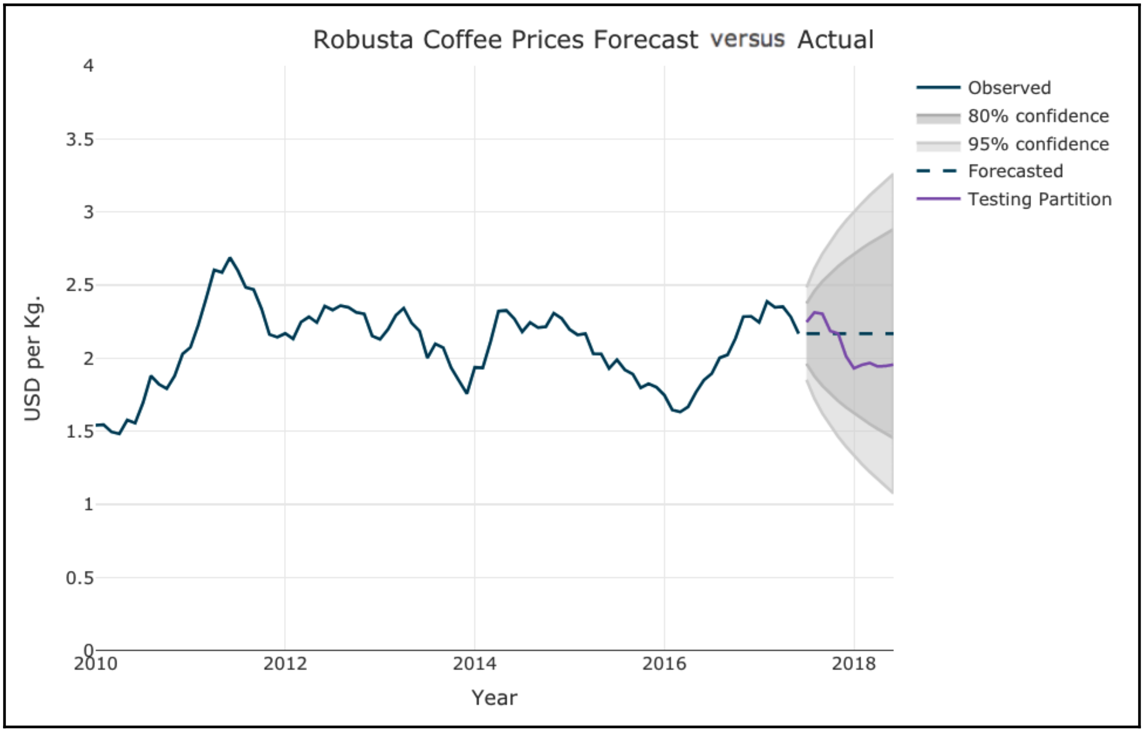

Forecasting the Robusta Coffee Monthly Prices
USD per Kg.
2.4
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
2015
2016
2017
2018
Actual
SMA - 1
SMA - 6
SMA - 12
SMA - 24
SMA - 36

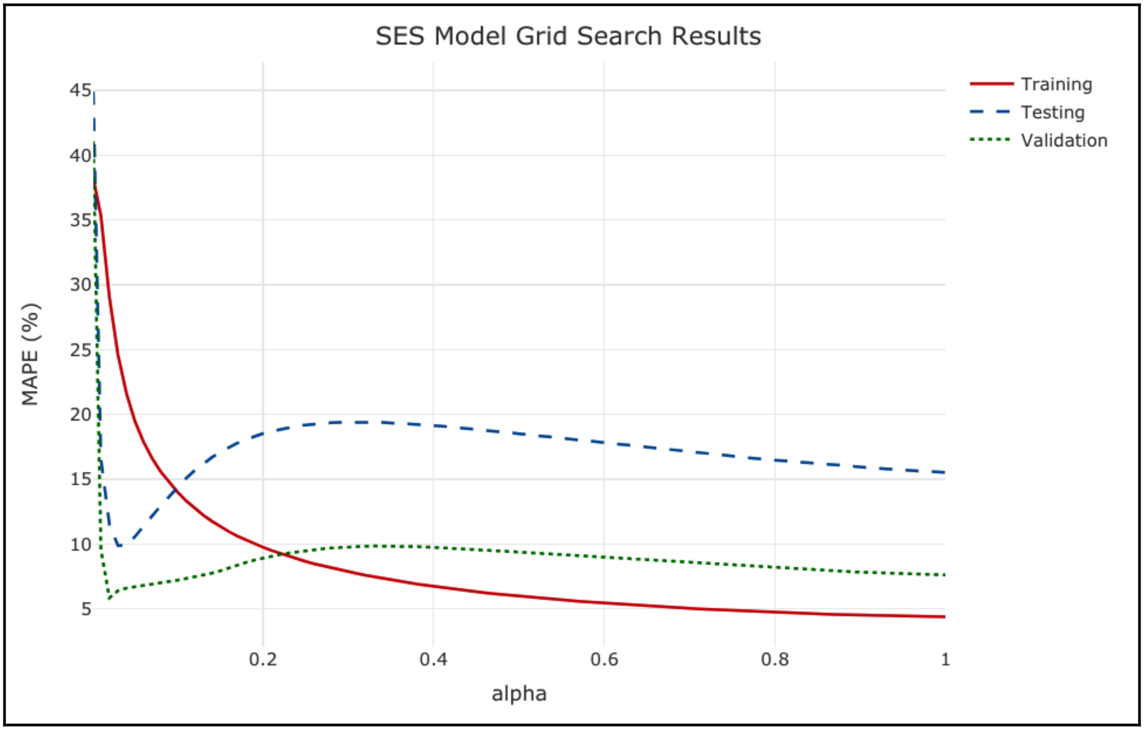


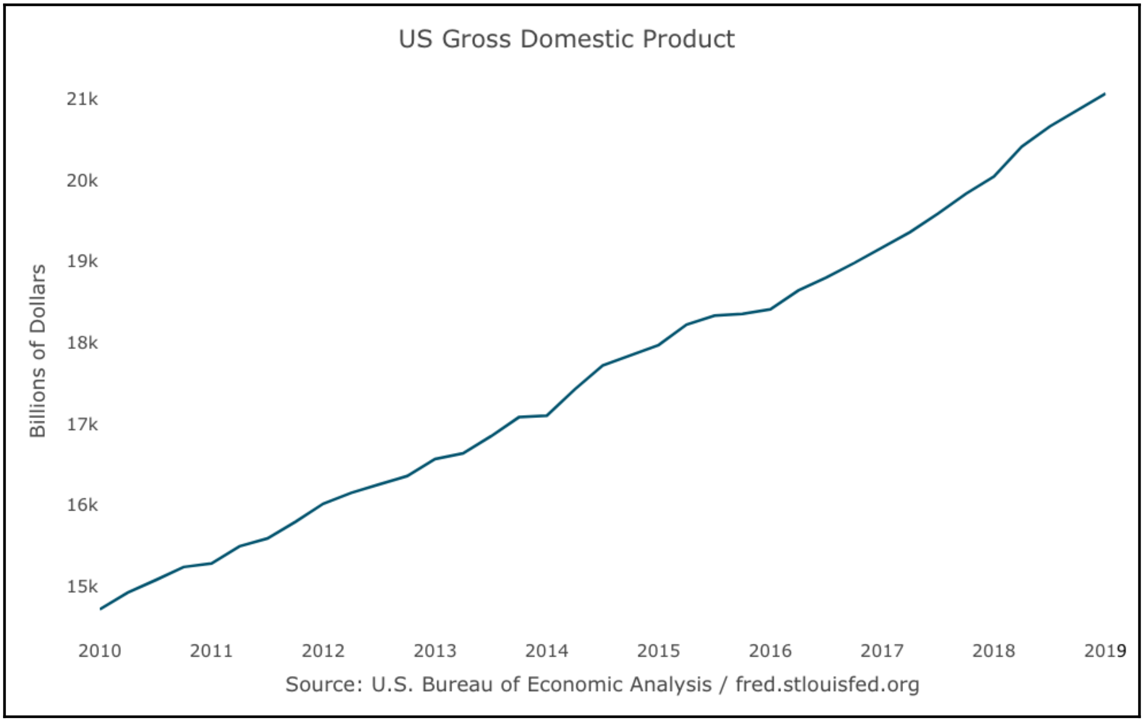
Forecasting the Monthly Consumption of Natural Gas in the US
Billion Cubic Feet
3200
3000
2800
2600
2400
2200
2000
Jan 2016
Jul 2016
Jan 2017
Jul 2017
Jan 2018
Jul 2018
Actual
WMA - Seasonal Lag
WMA - 12 (0.2/0.8)

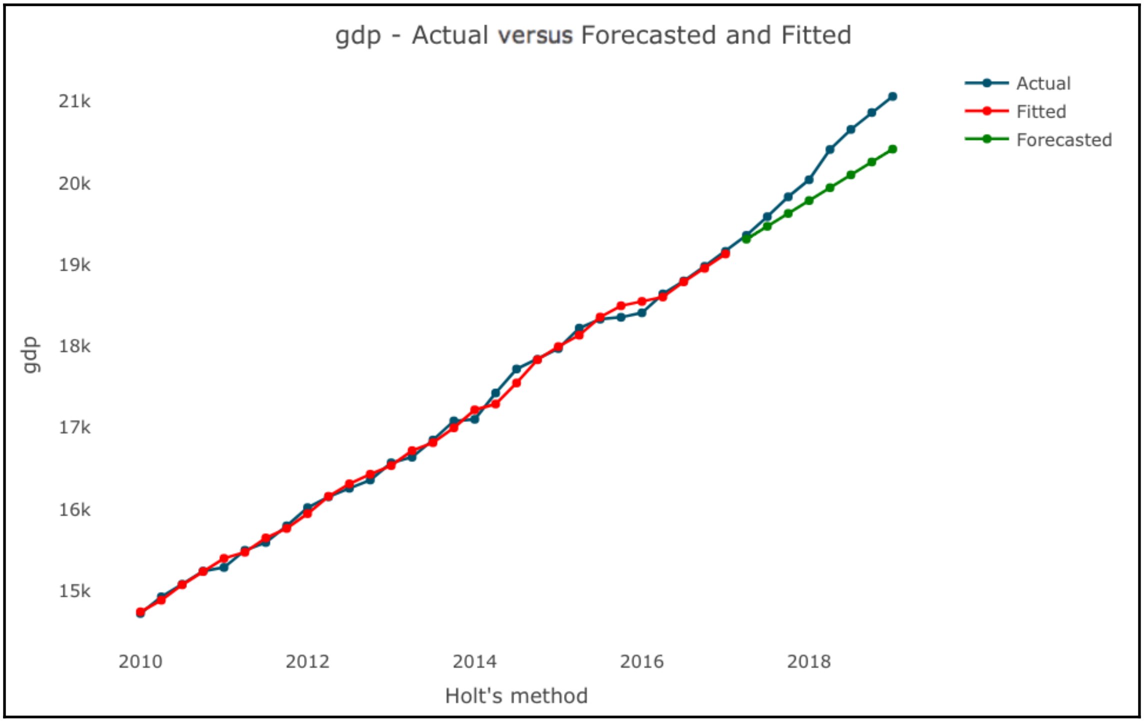

Decay Rate of the SES Weights
Weight
Index
1
0.8
0.6
0.4
0.2
0
2
4
6
8
10
12
14
alpha = 0.01
alpha = 0.2
alpha = 0.4
alpha = 0.6
alpha = 0.8
alpha = 1

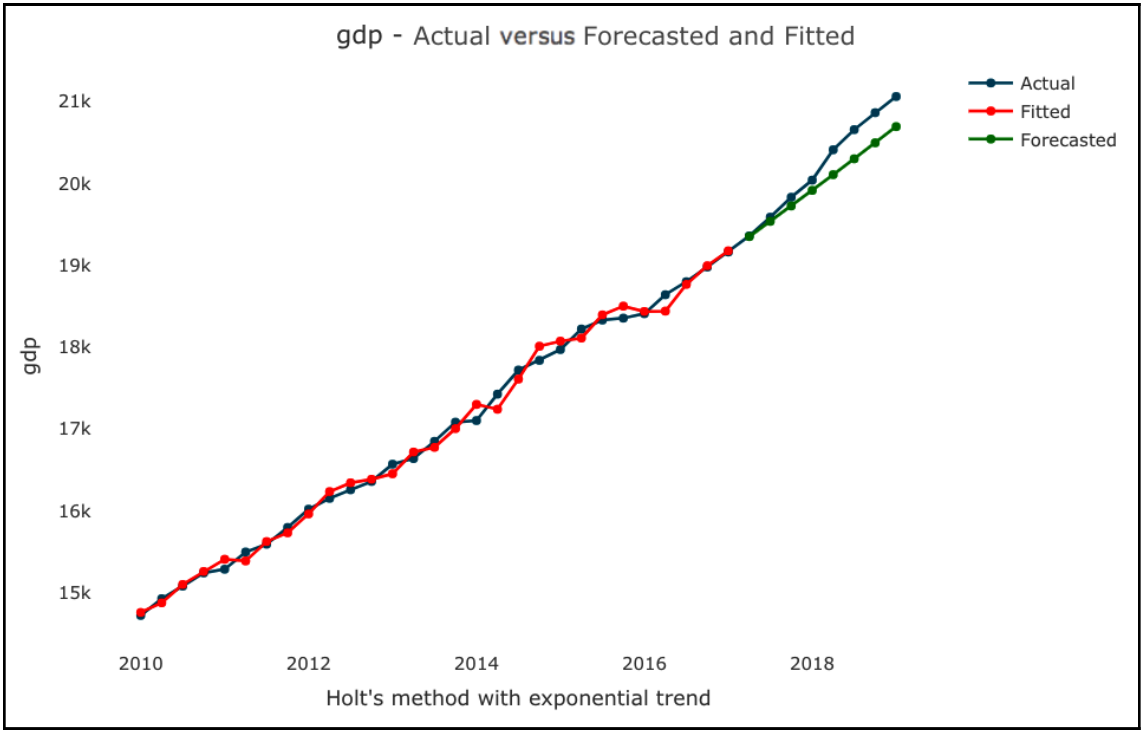

Robusta Coffee Prices Forecast versus Actual
Actual
Fitted
Forecasted
3
2.5
2
1.5
1
robusta
2010
2012
2014
2016
2018
Simple exponential smoothing

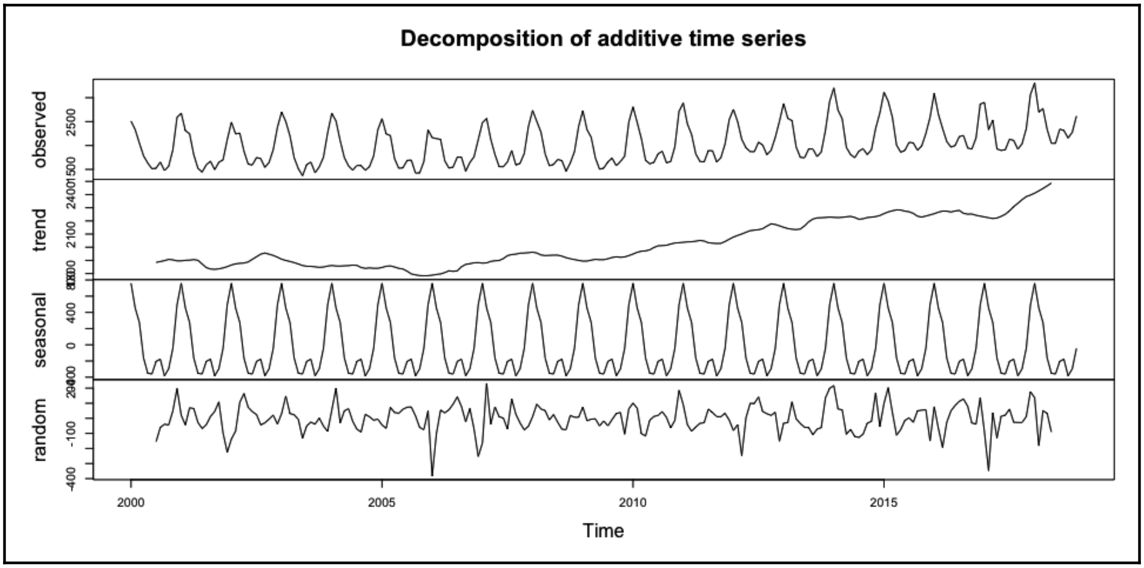

Robusta Coffee Prices Forecast versus Actual
4
3.5
3
2.5
2
1.5
1
0.5
0
USD per Kg.
2010
2012
2014
2016
2018
Year
Observed
80% confidence
95% confidence
Forecasted
Testing Partition

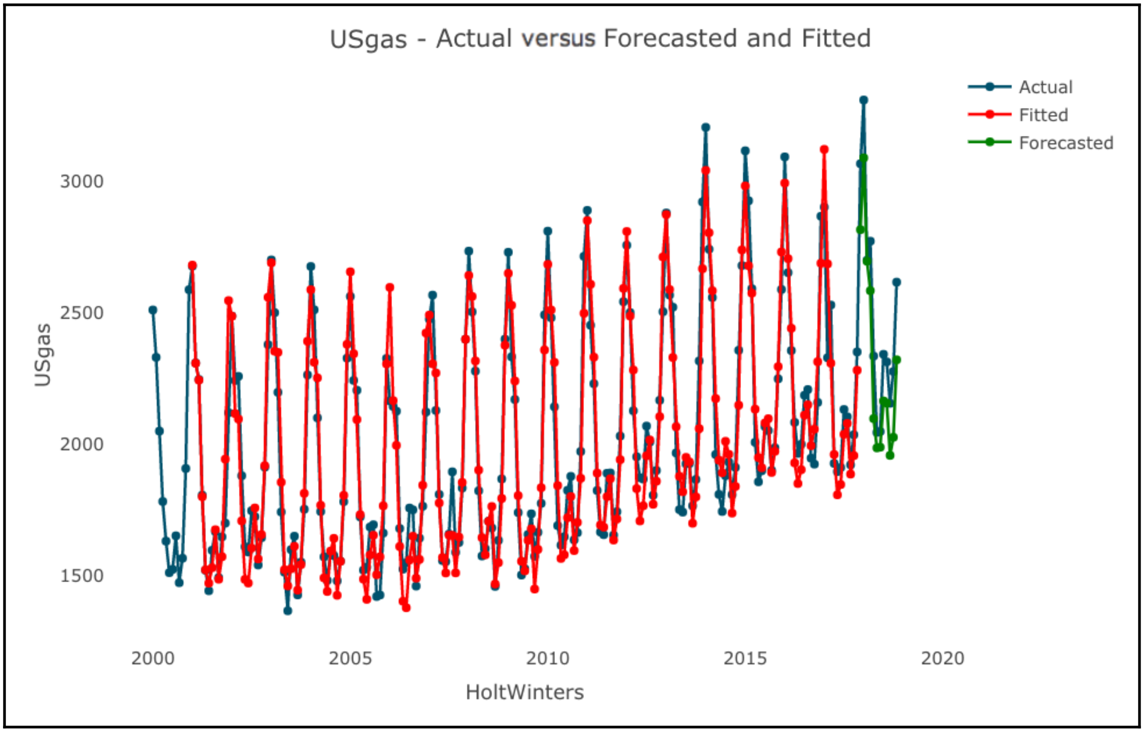

SES Model Grid Search Results
Training
Testing
Validation
MAPE (%)
45
40
35
30
25
20
15
10
5
0.2
0.4
0.6
0.8
1
alpha

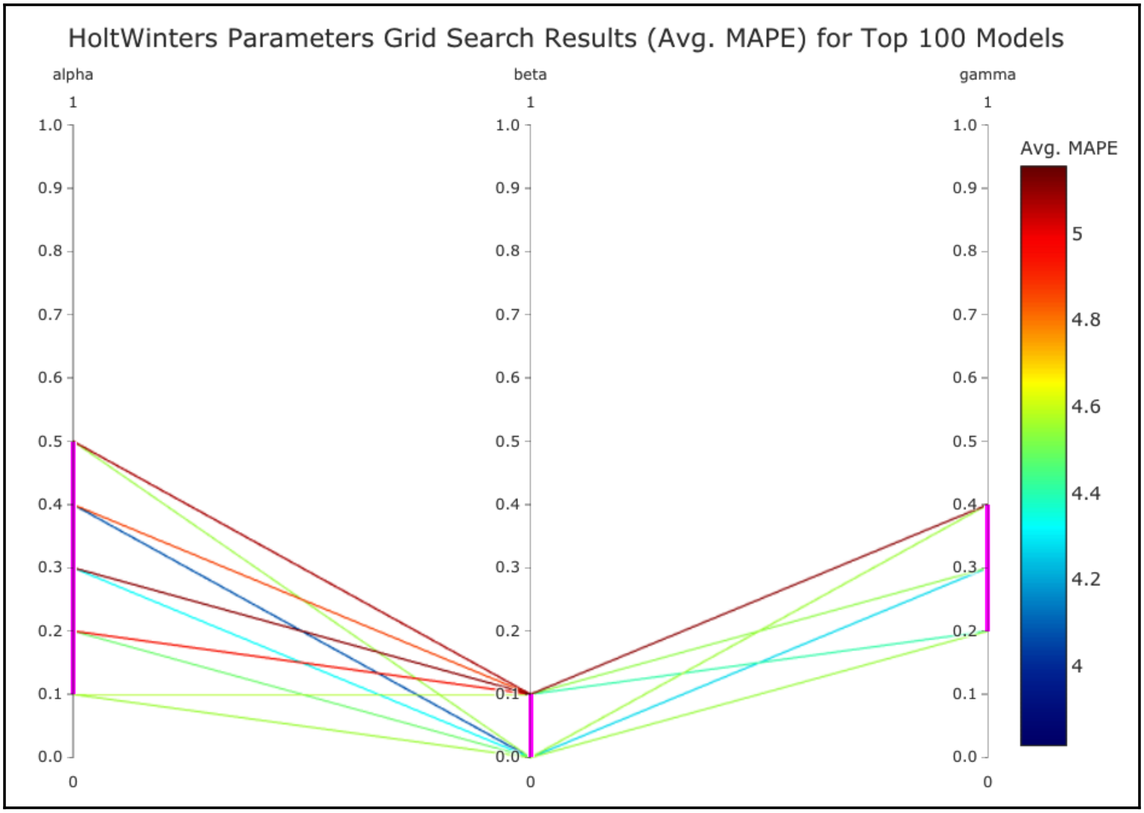

US Gross Domestic Product
Billions of Dollars
21k
20k
19k
18k
17k
16k
15k
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Source: U.S. Bureau of Economic Analysis / fred.stlouisfed.org


gdp - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
21k
20k
19k
18k
17k
16k
15k
gdp
2010
2012
2014
2016
2018
Holt's method

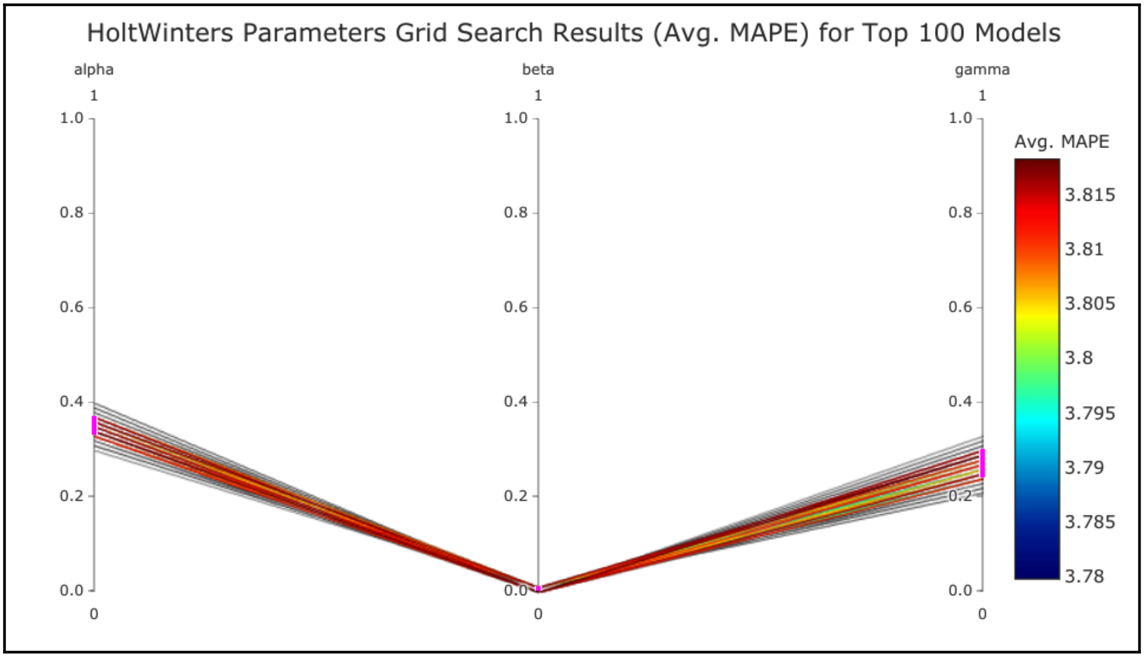

gdp - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
21k
20k
19k
18k
17k
16k
15k
gdp
2010
2012
2014
2016
2018
Holt's method with exponential trend


Decomposition of additive time series
observed
trend
seasonal
random
2000
2005
2010
2015
Time

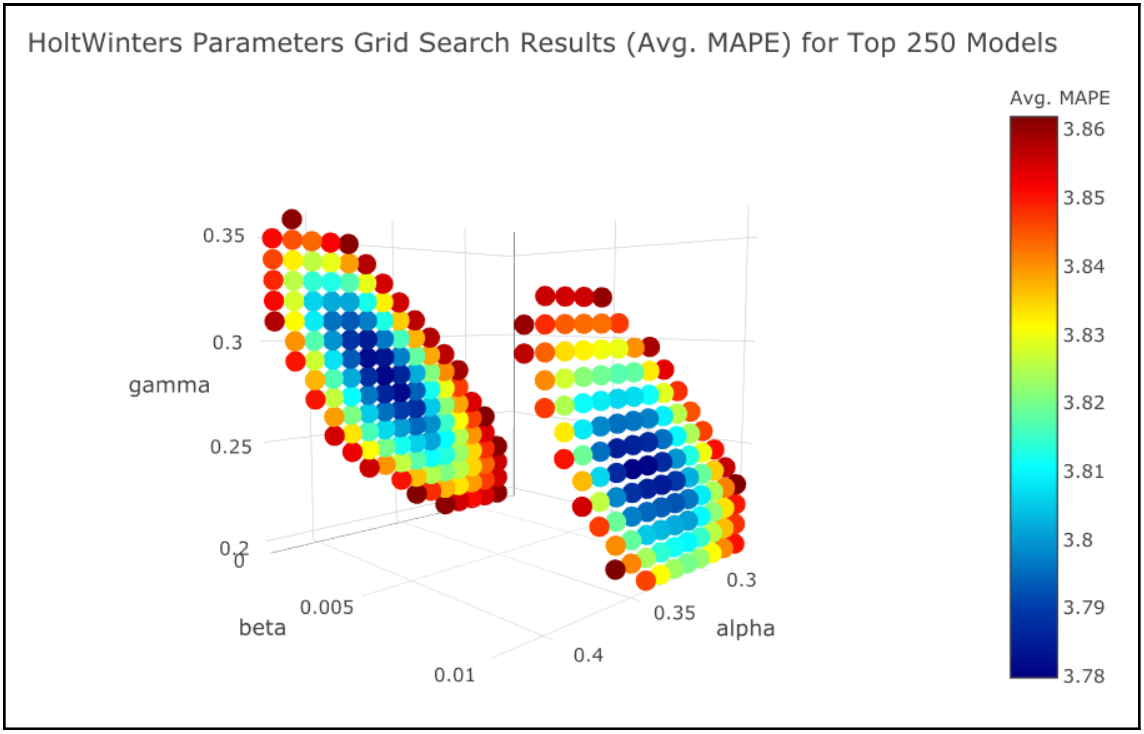

USgas - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
3000
2500
2000
1500
USgas
2000
2005
2010
2015
2020
HoltWinters

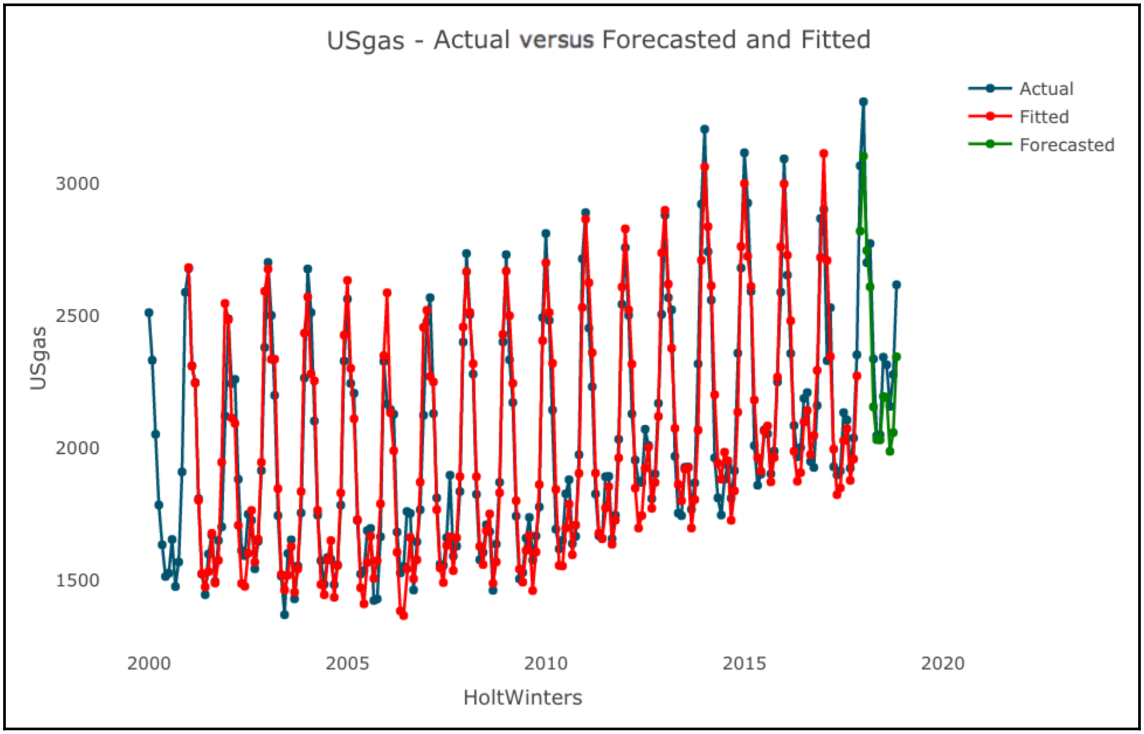


HoltWinters Parameters Grid Search Results (Avg. MAPE) for Top 100 Models
alpha
beta
gamma
1
0
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Avg. MAPE
5
4.8
4.6
4.4
4.2
4

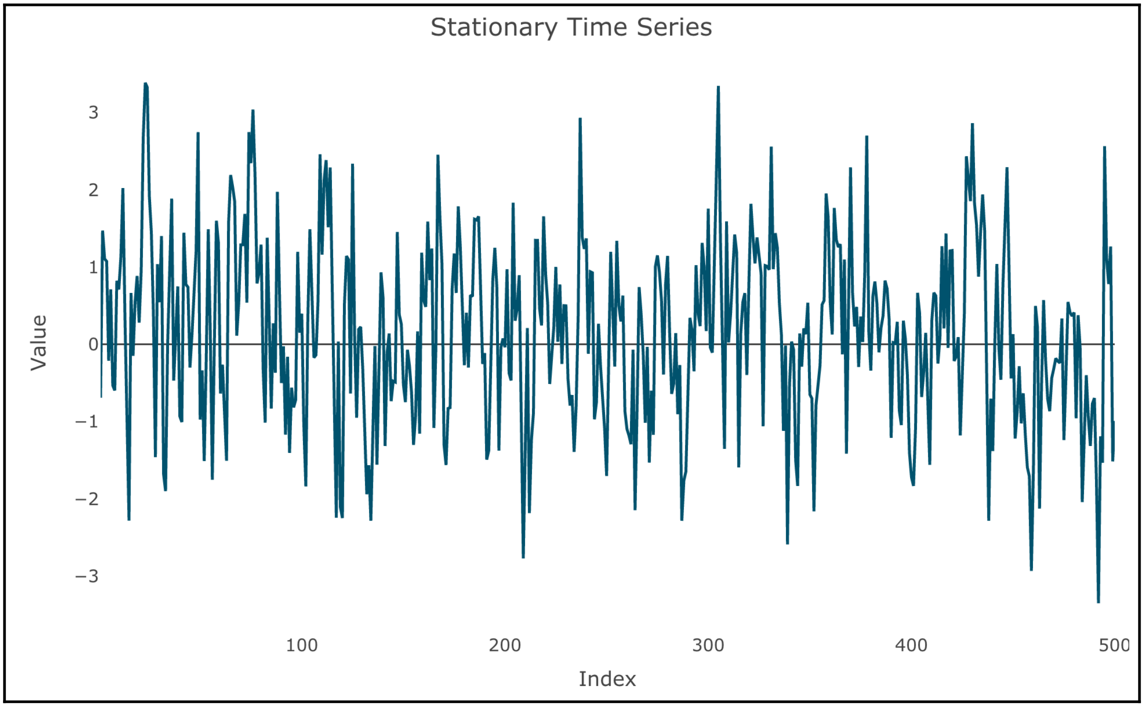


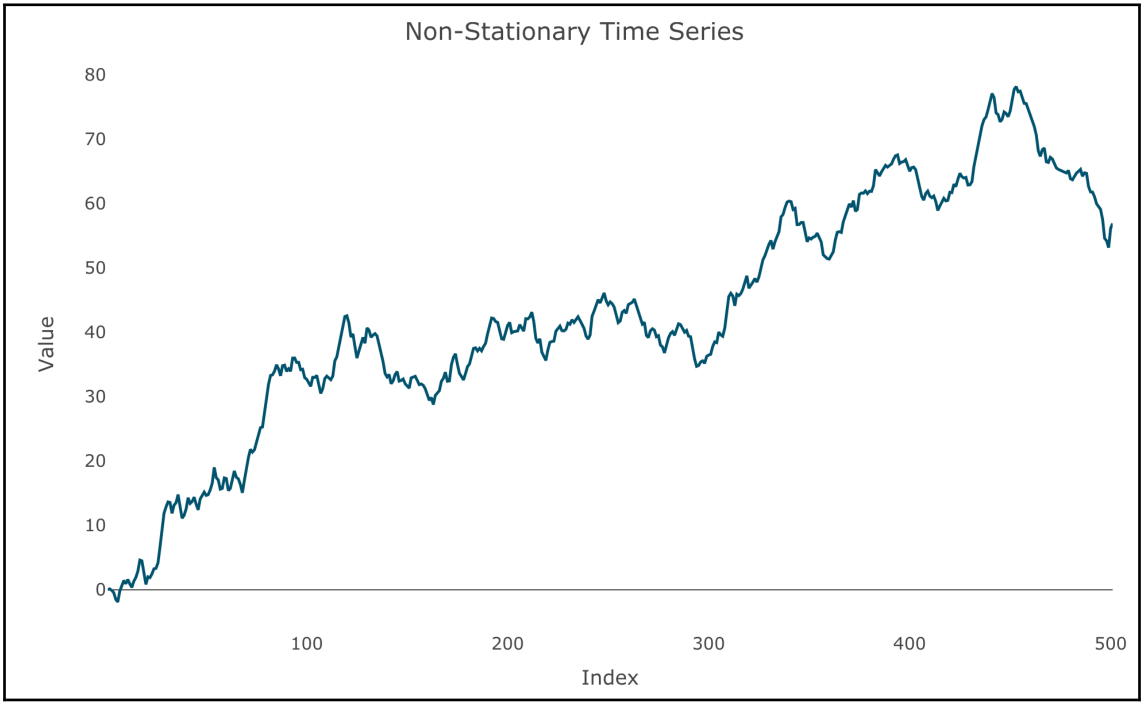
HoltWinters Parameters Grid Search Results (Avg. MAPE) for Top 100 Models
alpha
beta
gamma
1
1.0
0.8
0.6
0.4
0.2
0.0
0
Avg. MAPE
3.815
3.81
3.805
3.8
3.795
3.79
3.785
3.78

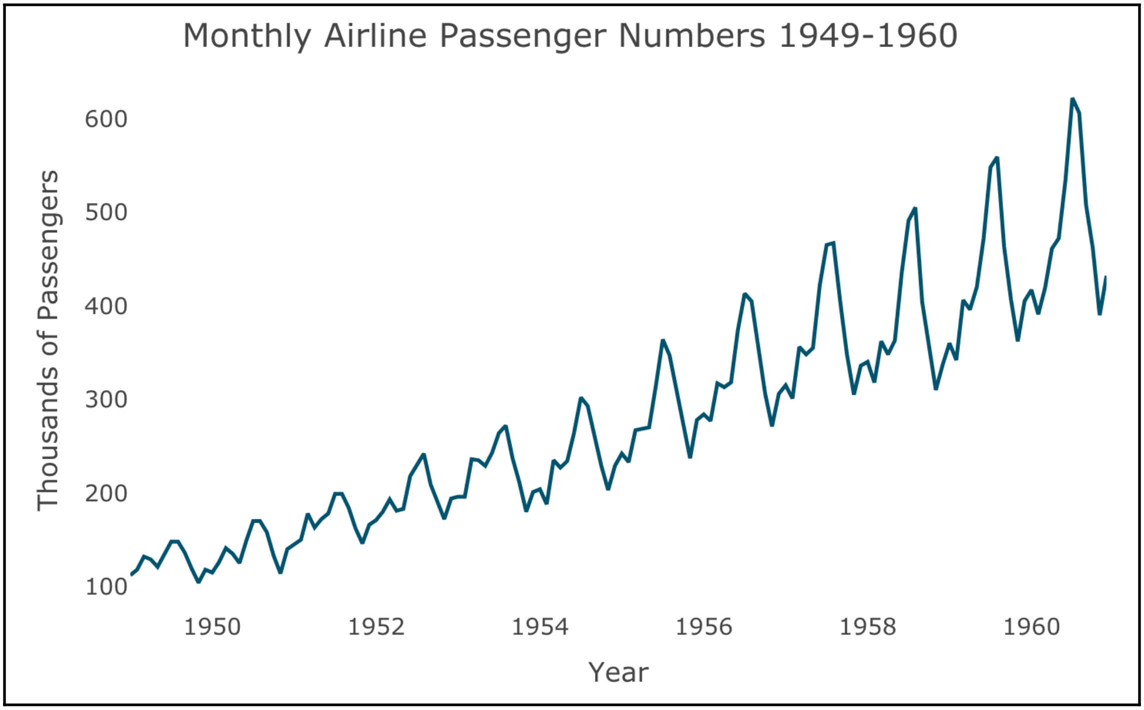

HoltWinters Parameters Grid Search Results (Avg. MAPE) for Top 250 Models
Avg. MAPE
3.86
3.85
3.84
3.83
3.82
3.81
3.8
3.79
3.78
gamma
0.35
0.3
0.25
0.2
0
beta
0.005
0.01
alpha
0.3
0.35
0.4

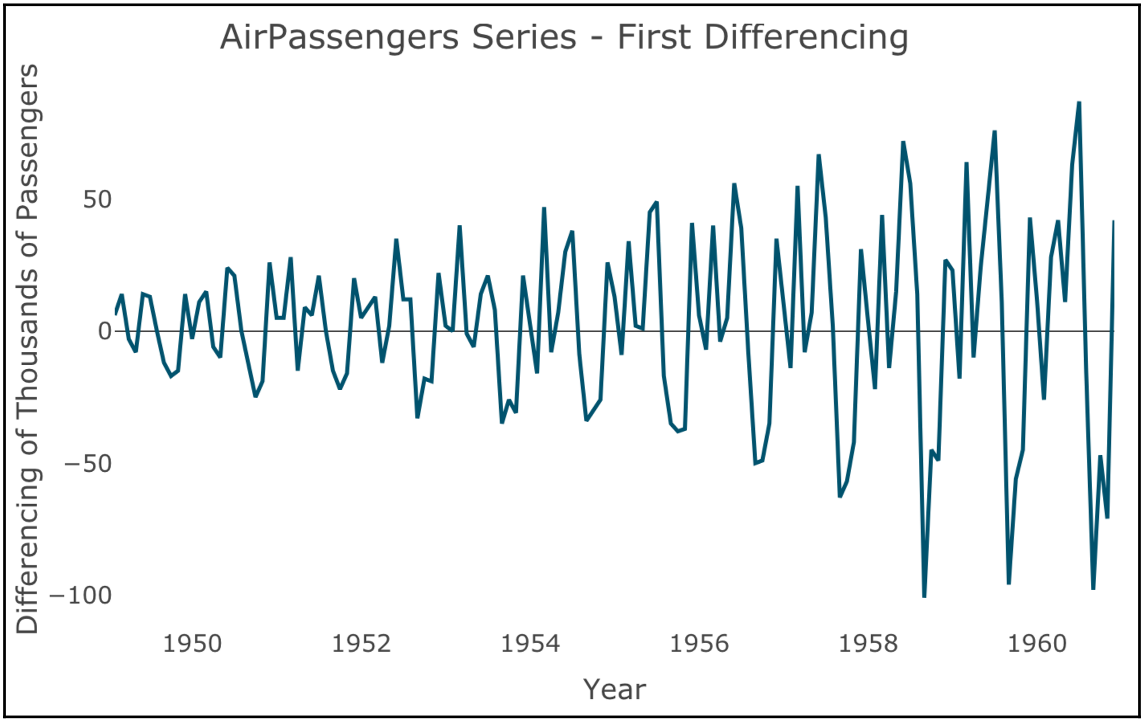

USgas - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
USgas
3000
2500
2000
1500
2000
2005
2010
2015
2020
HoltWinters

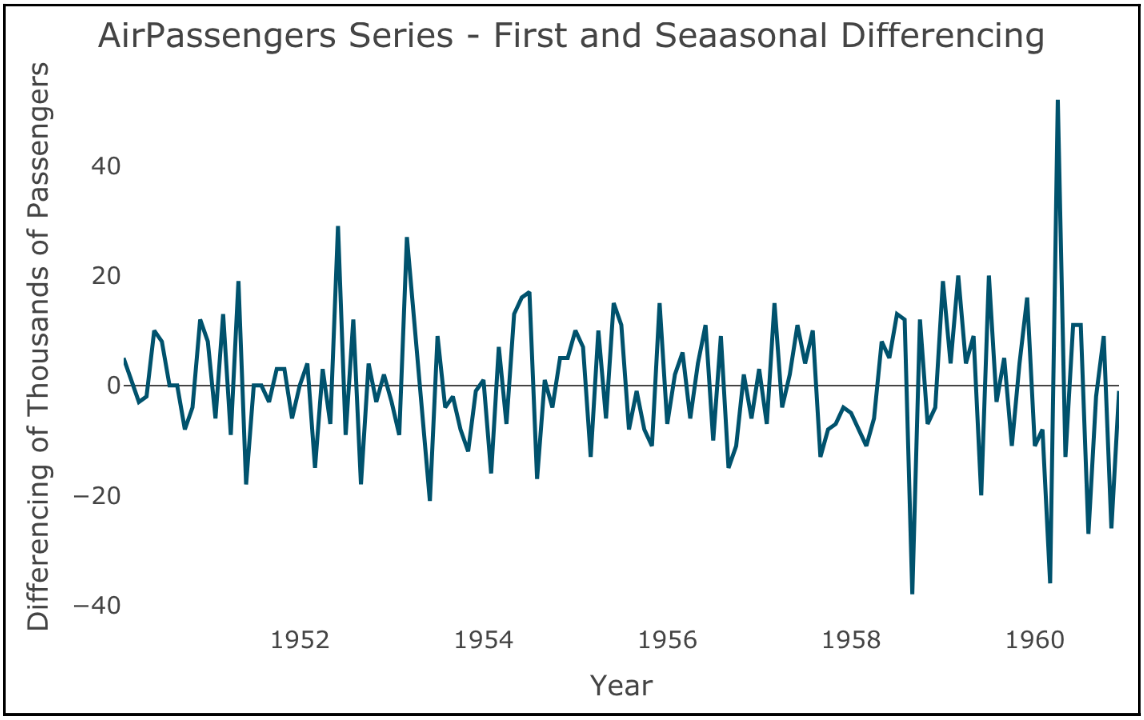
Chapter 11: Forecasting with ARIMA Models

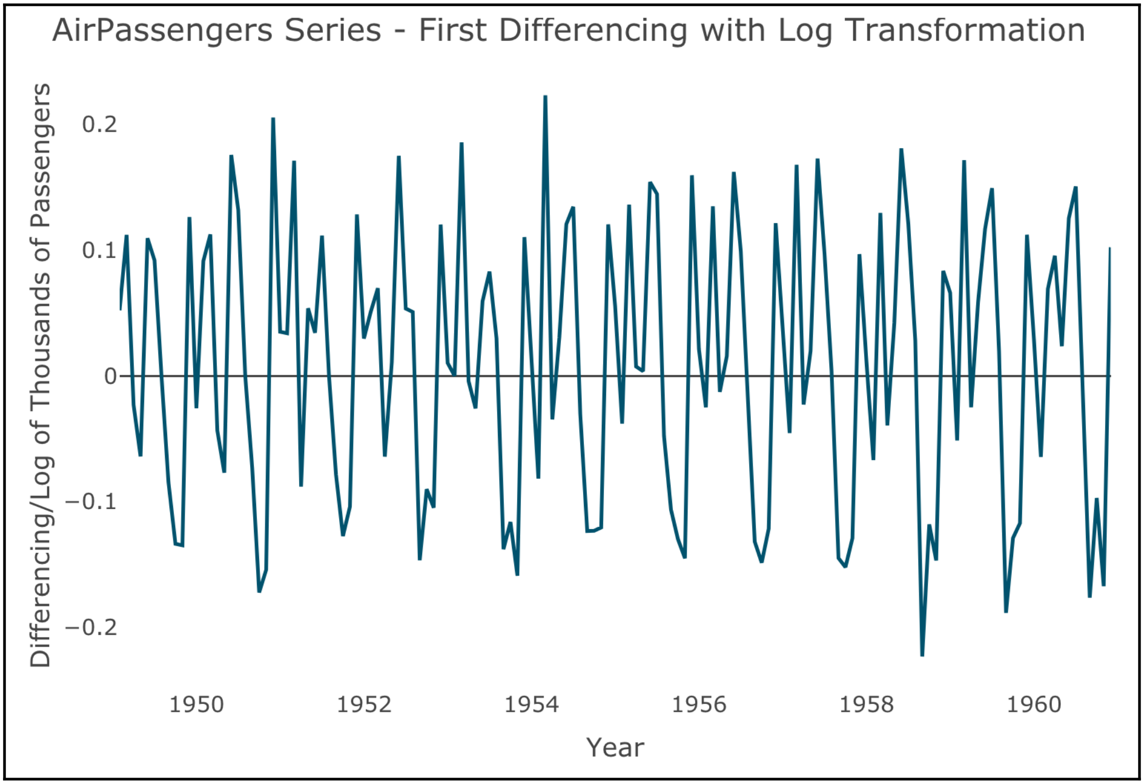


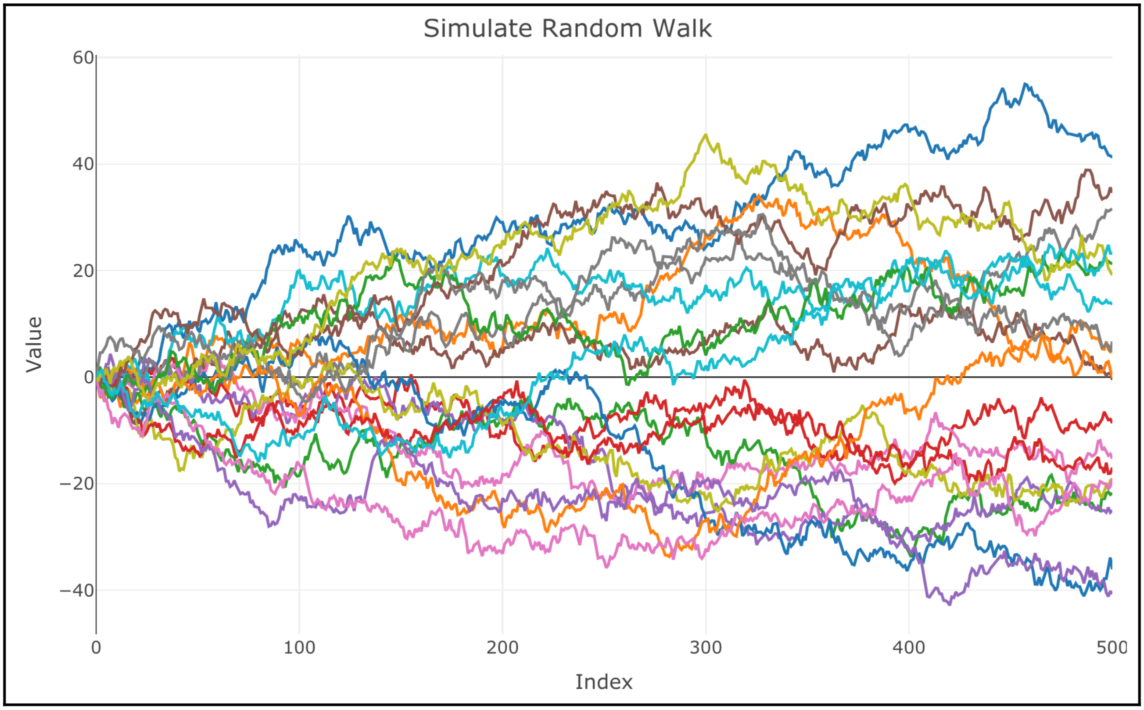

Non-Stationary Time Series
Value
80
70
60
50
40
30
20
10
0
100
200
300
400
500
Index


Monthly Airline Passenger Numbers 1949-1960
600
500
400
300
200
100
Thousands of Passengers
1950
1952
1954
1956
1958
1960
Year

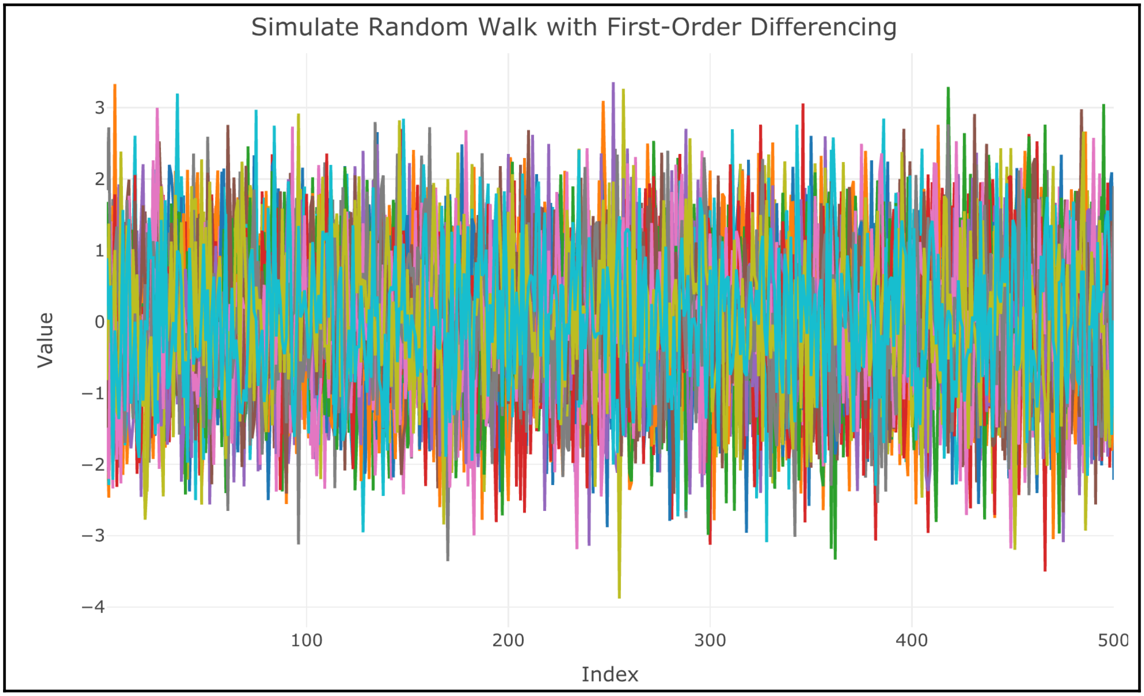

AirPassengers Series - First Differencing
Differencing of Thousands of Passengers
50
0
−50
−100
1950
1952
1954
1956
1958
1960
Year

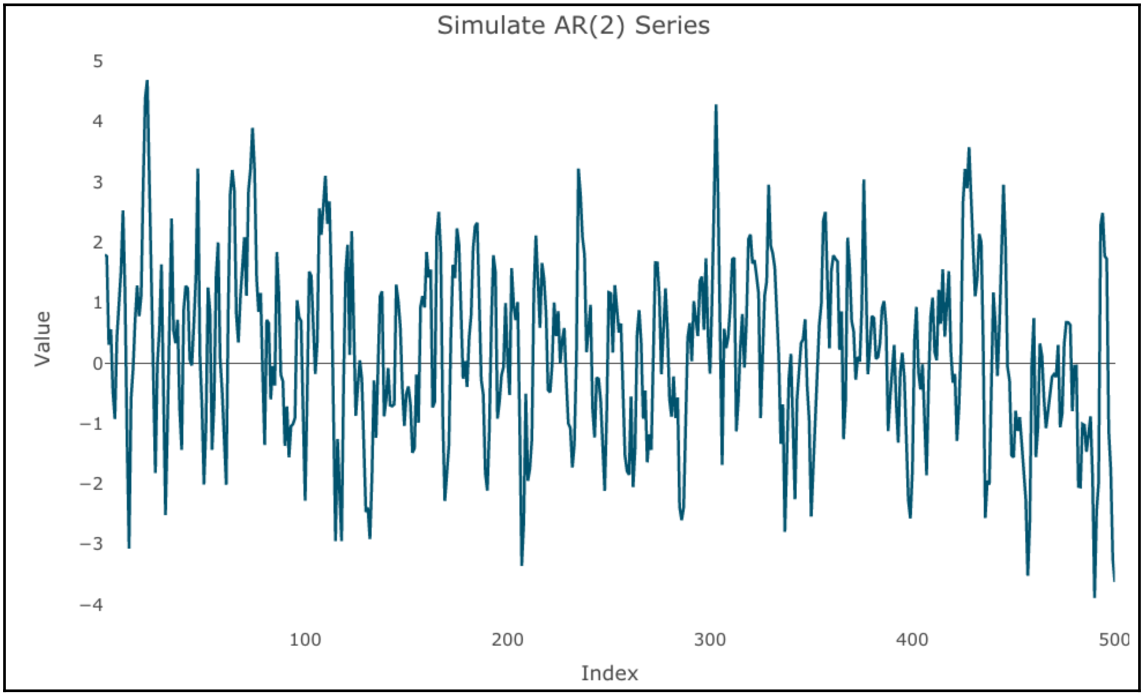

AirPassengers Series - First and Seaasonal Differencing
Differencing of Thousands of Passengers
40
20
0
−20
−40
1952
1954
1956
1958
1960
Year

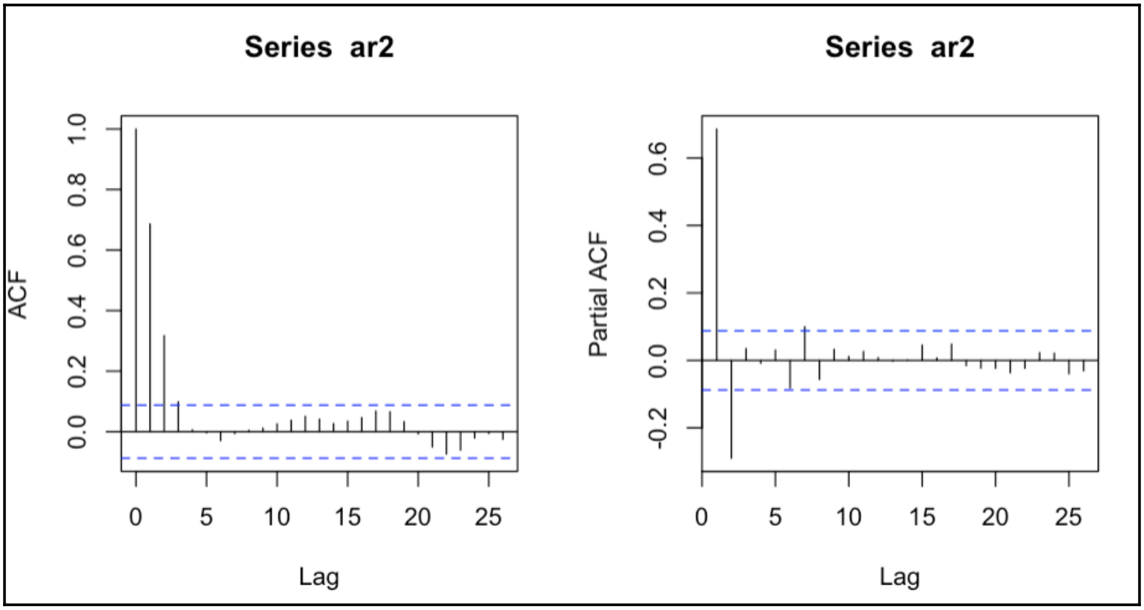

AirPassengers Series - First Differencing with Log Transformation
Differencing/Log of Thousands of Passengers
0.2
0.1
0
−0.1
−0.2
1950
1952
1954
1956
1958
1960
Year

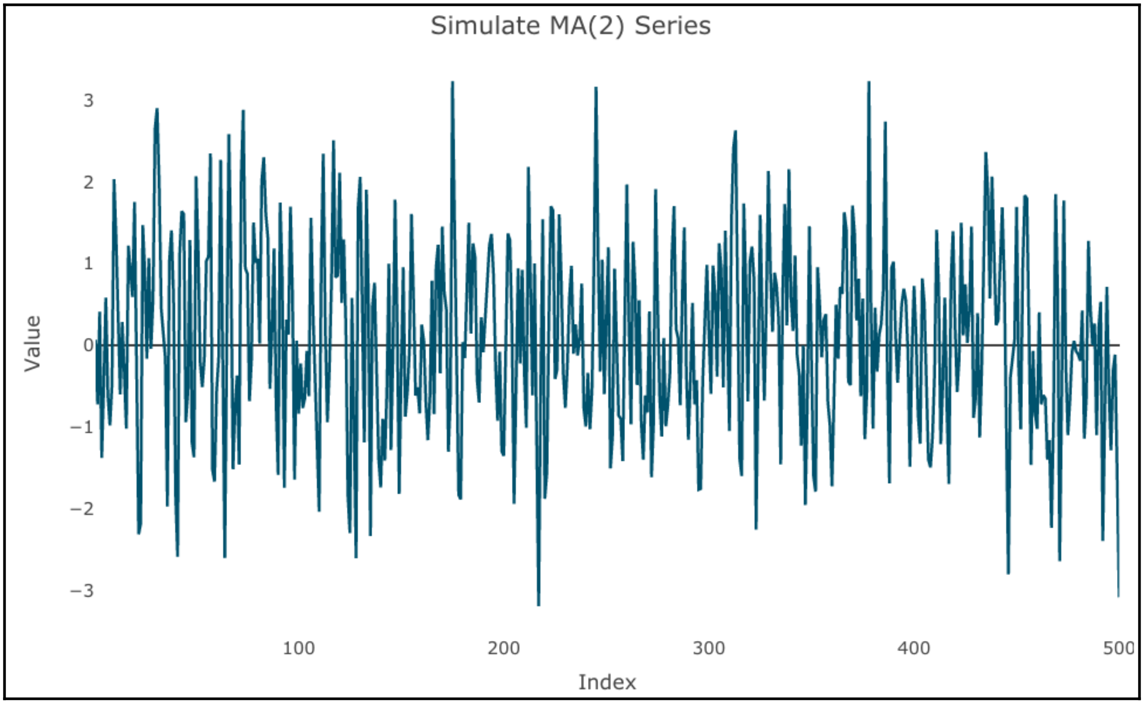


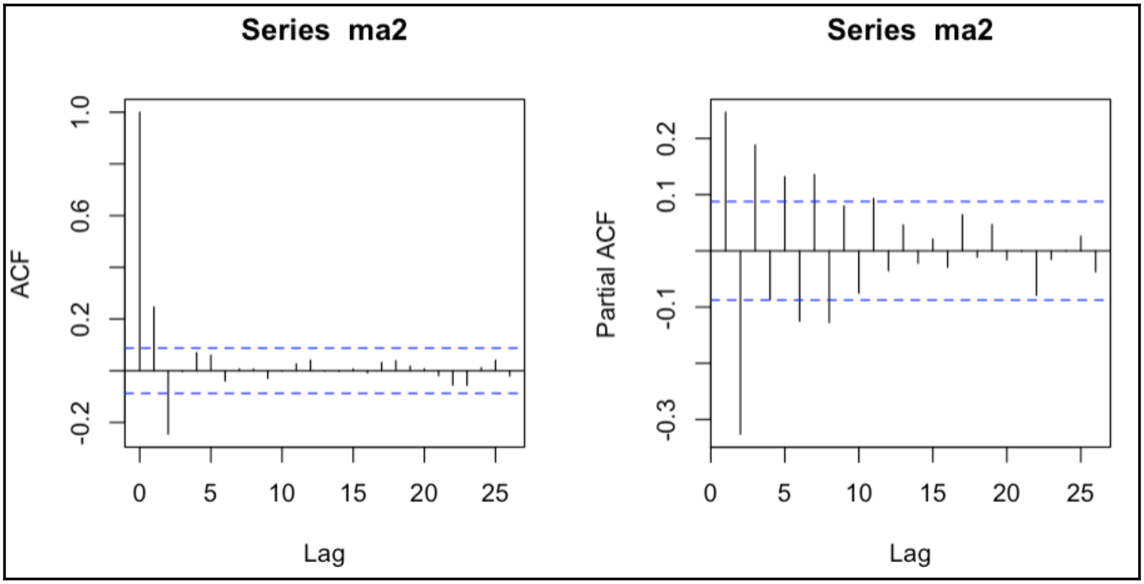
Simulate Random Walk
Value
Index
60
40
20
0
−20
−40
0
100
200
300
400
500

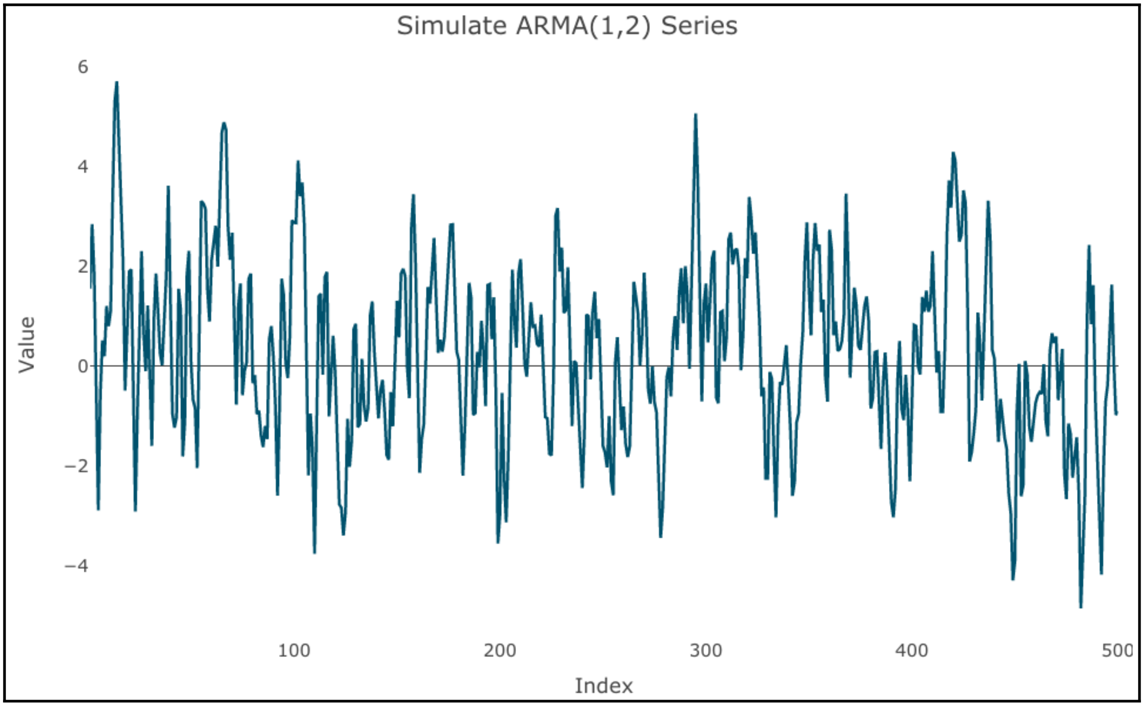

Simulate Random Walk with First-Order Differencing
Value
3
2
1
0
−1
−2
−3
−4
100
200
300
400
500
Index

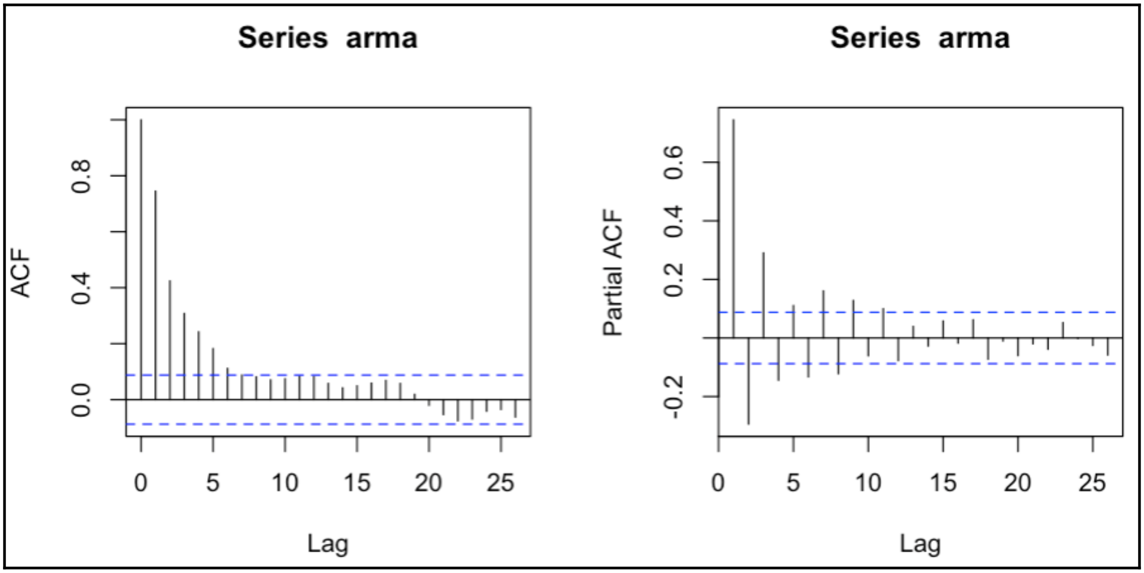

Simulate AR(2) Series
Value
Index

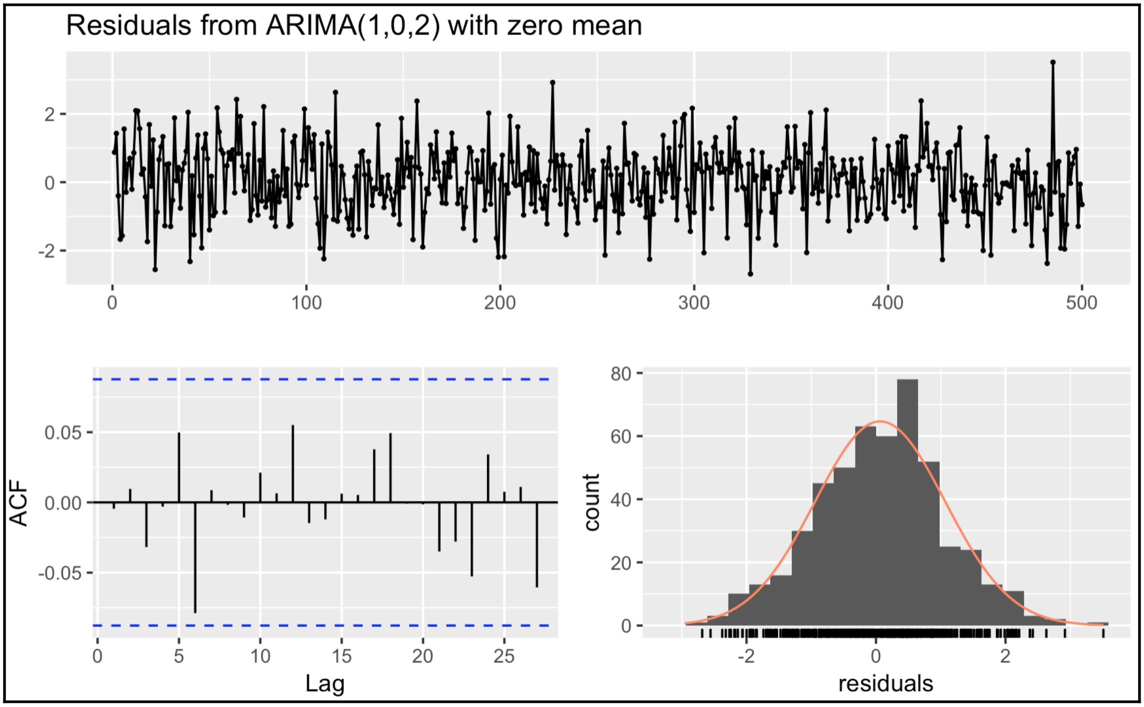

Series ar2
ACF
Lag
Series ar2
Partial ACF
Lag

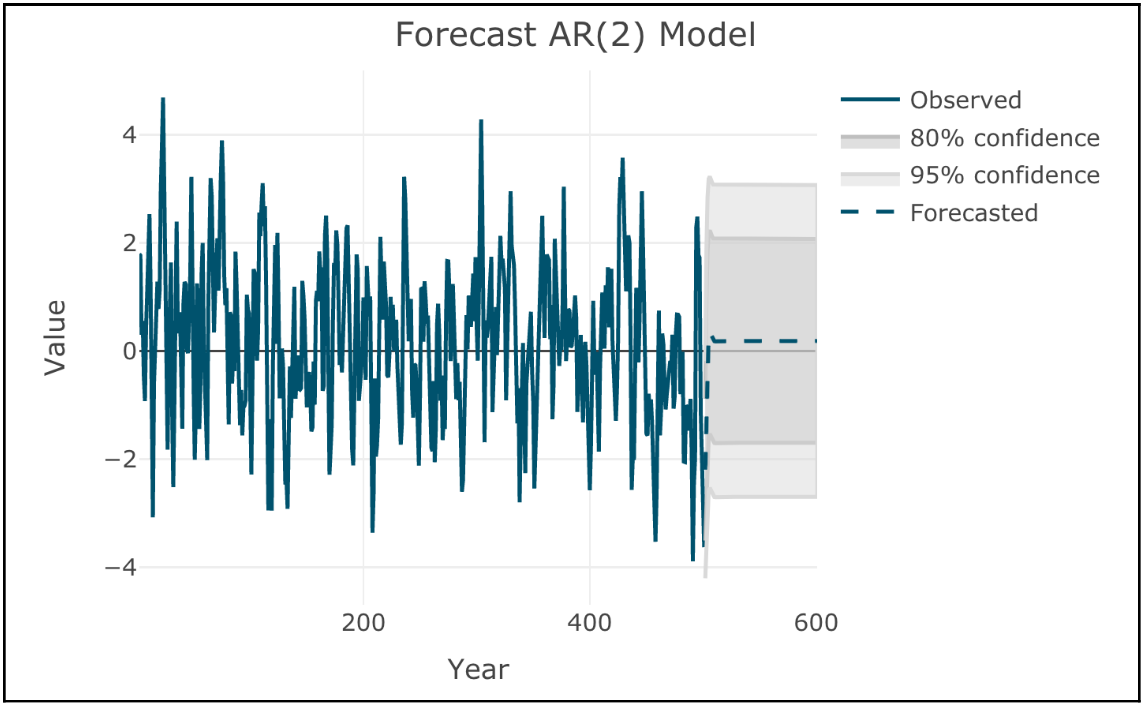

Simulate MA(2) Series
Value
3
2
1
0
−1
−2
−3
100
200
300
400
500
Index

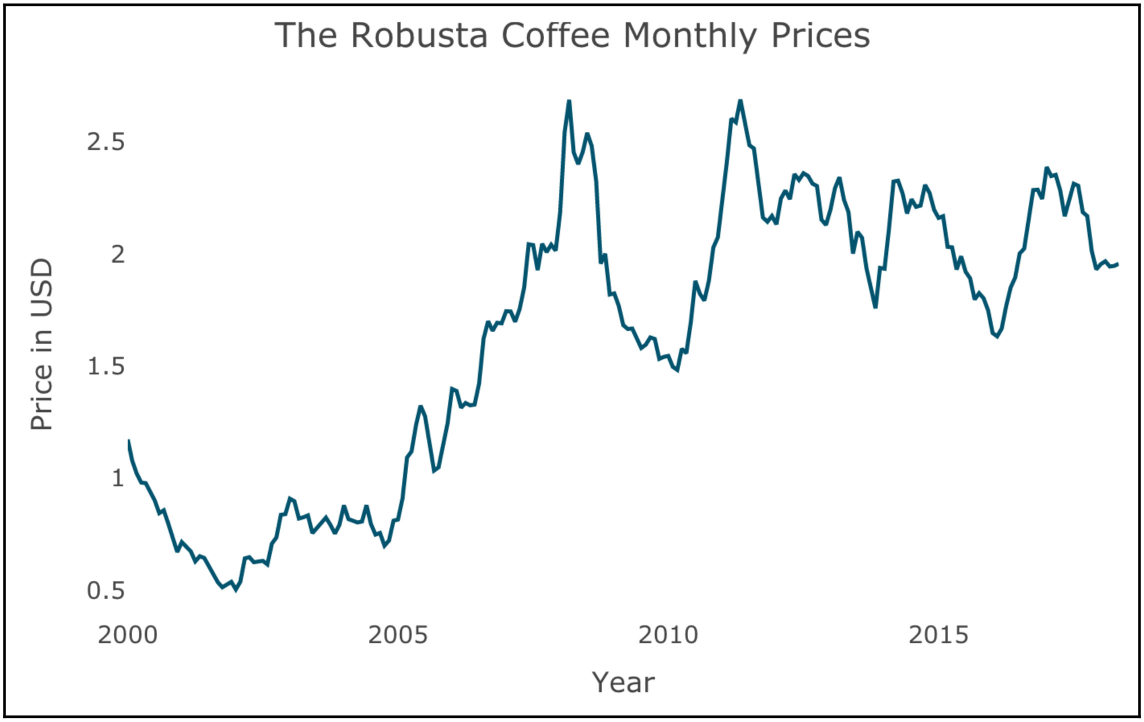

Series ma2
ACF
1.0
0.6
0.2
-0.2
0
5
10
15
20
25
Lag
Series ma2
Partial ACF
0.2
0.1
-0.1
-0.3
0
5
10
15
20
25
Lag

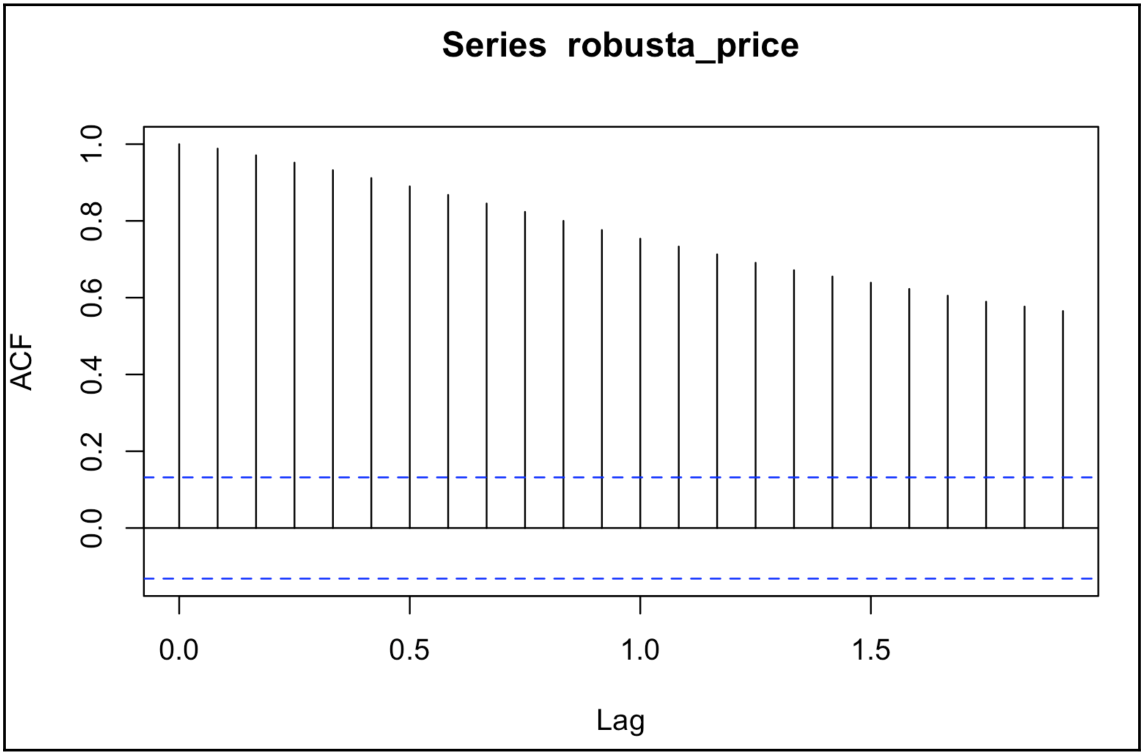

Simulate ARMA(1,2) Series
Value
6
4
2
0
−2
−4
100
200
300
400
500
Index

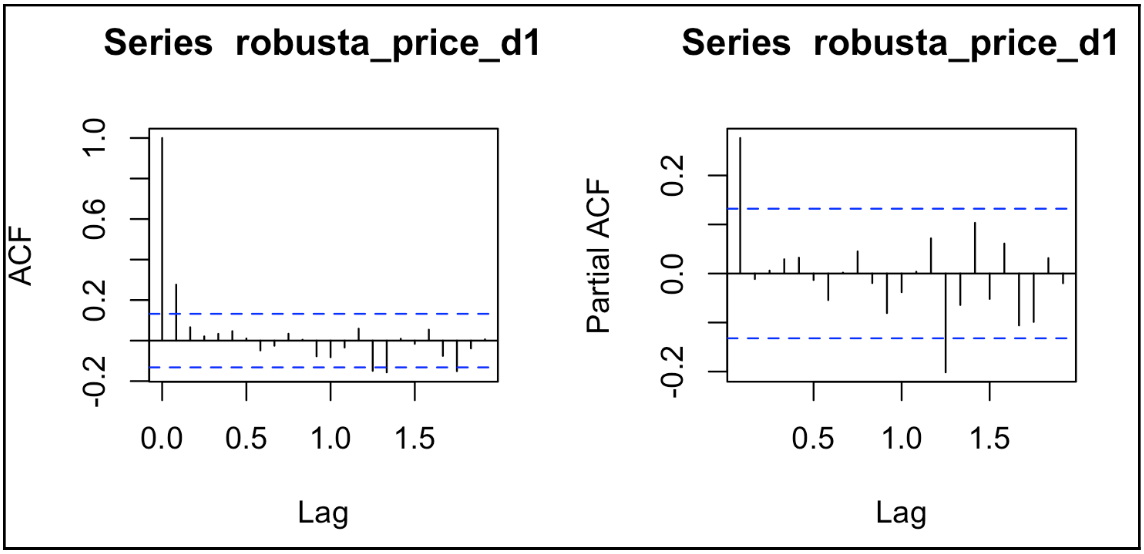

Series arma
ACF
0.0
0.4
0.8
0
5
10
15
20
25
Lag
Series arma
Partial ACF
-0.2
0.2
0.4
0.6
0
5
10
15
20
25
Lag

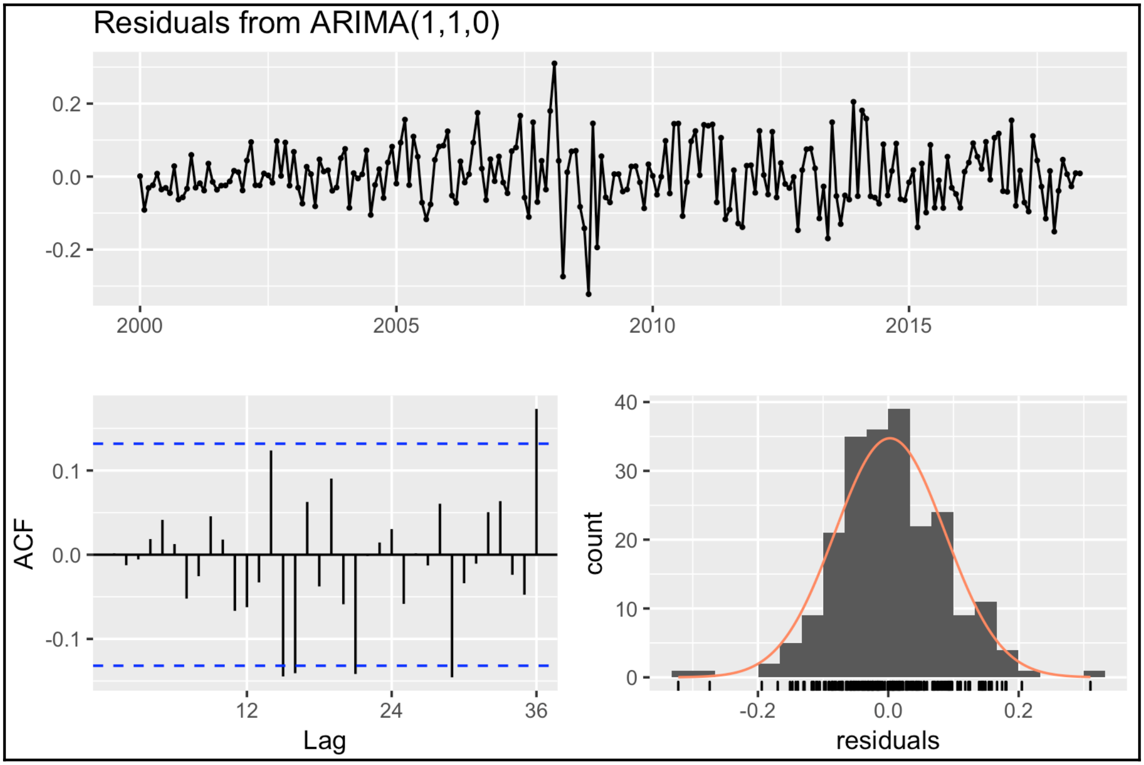

Residuals from ARIMA(1,0,2) with zero mean
2
0
-2
0
100
200
300
400
500
0.05
0.00
-0.05
ACF
0
5
10
15
20
25
Lag
80
60
40
20
0
count
-2
0
2
residuals

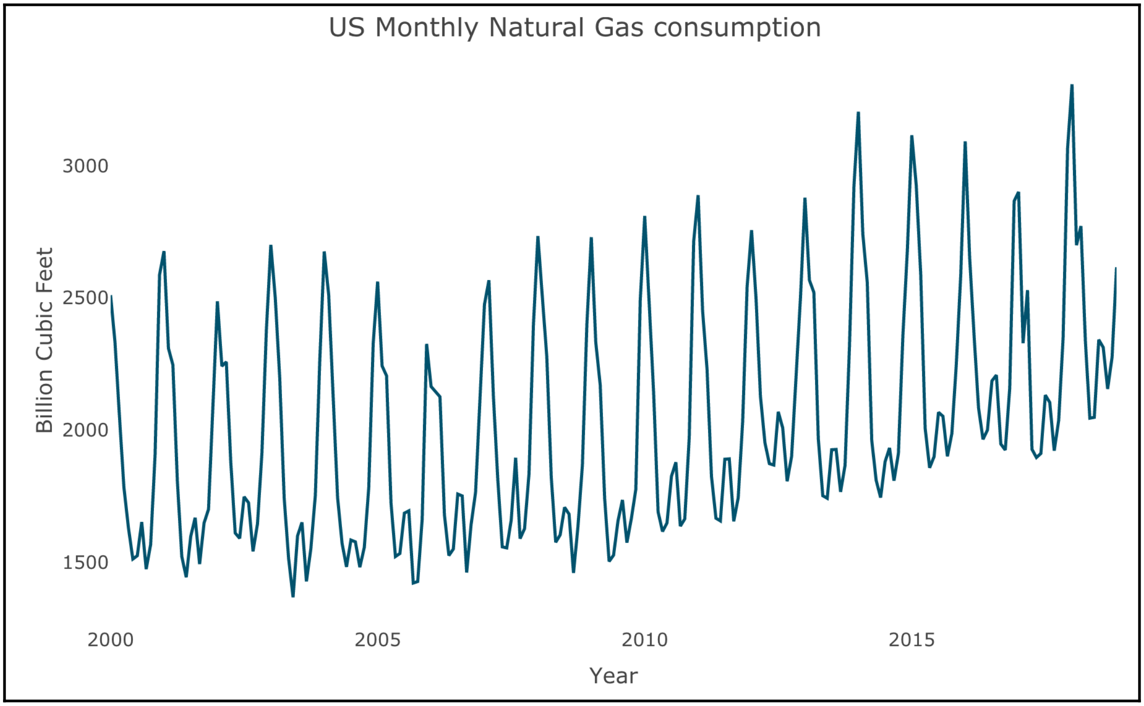

Forecast AR(2) Model
Value
4
2
0
−2
−4
200
400
600
Year
Observed
80% confidence
95% confidence
Forecasted

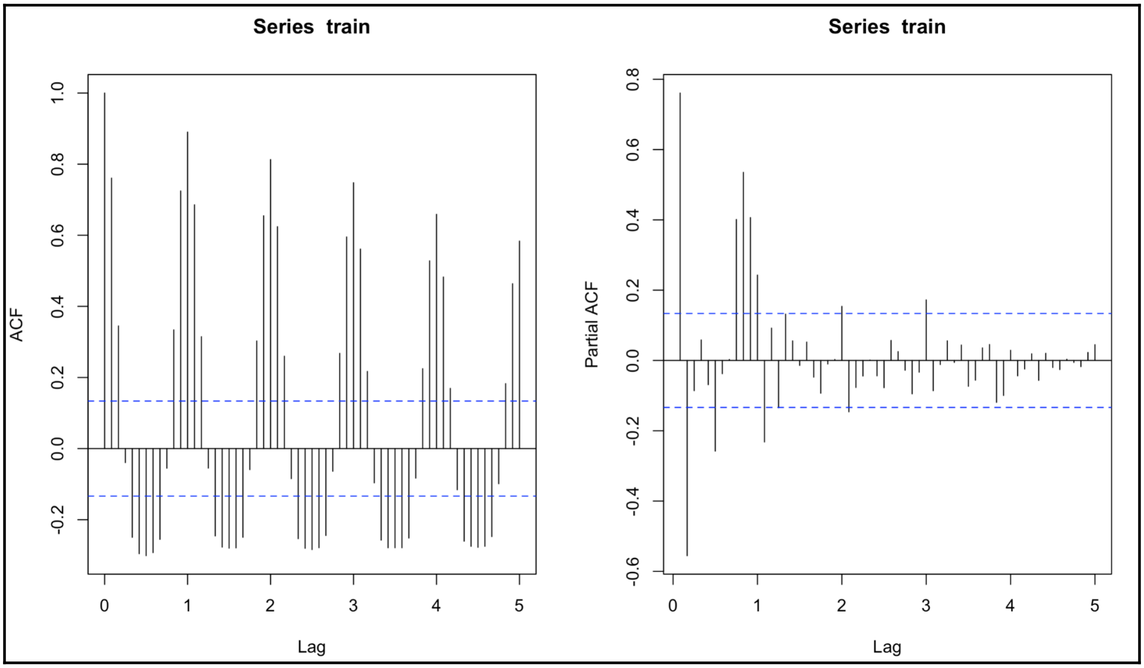

The Robusta Coffee Monthly Prices
Price in USD
2.5
2
1.5
1
0.5
2000
2005
2010
2015
Year

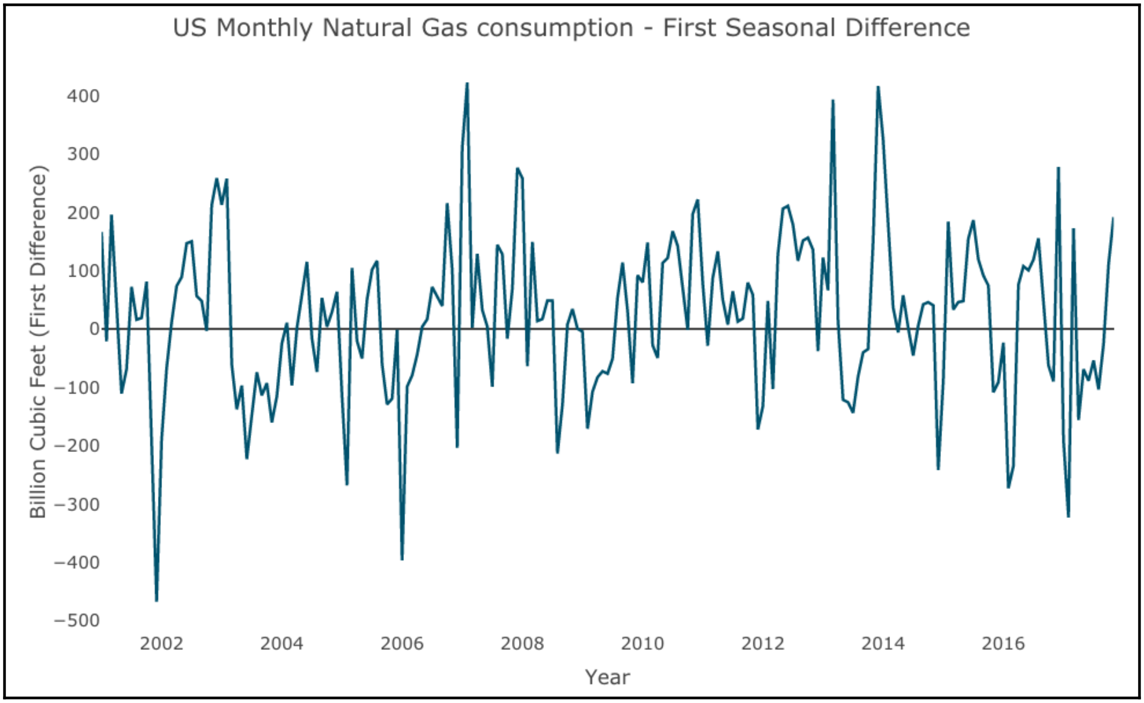

Series robusta_price
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
1.5
Lag

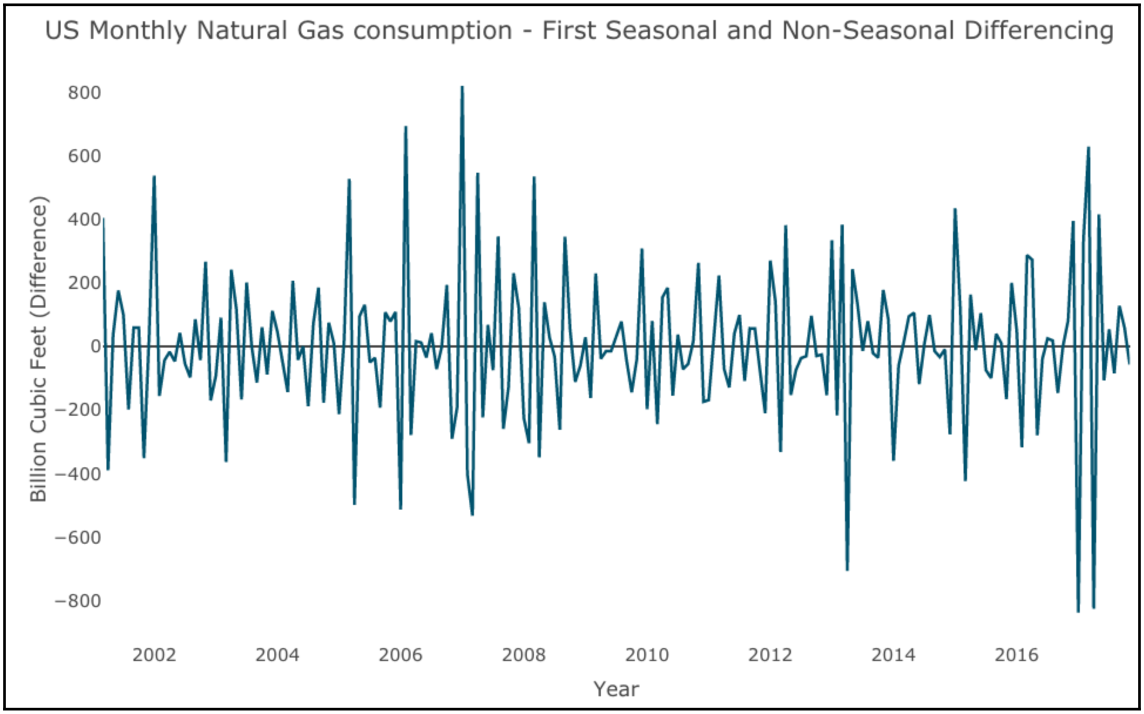

Series robusta_price_d1
ACF
1.0
0.6
0.2
-0.2
0.0
0.5
1.0
1.5
Lag
Series robusta_price_d1
Partial ACF
0.2
0.0
-0.2
0.5
1.0
1.5
Lag

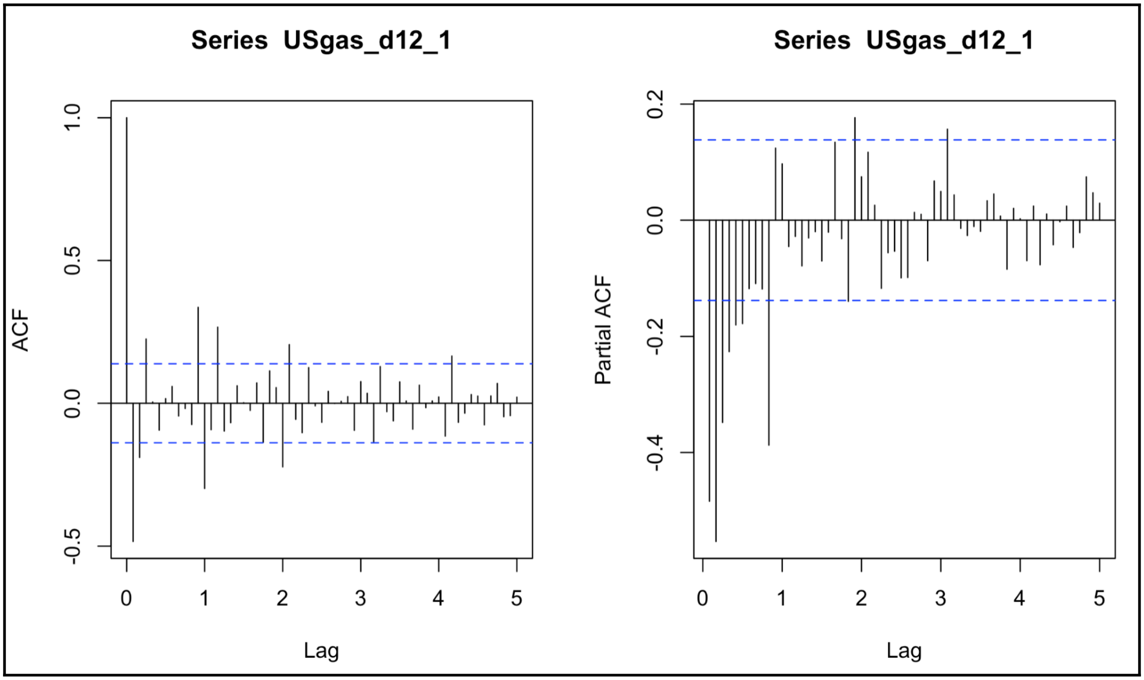

Residuals from ARIMA(1,1,0)
0.2
0.0
-0.2
2000
2005
2010
2015
ACF
0.1
0.0
-0.1
12
24
36
Lag
count
40
30
20
10
0
-0.2
0.0
0.2
residuals

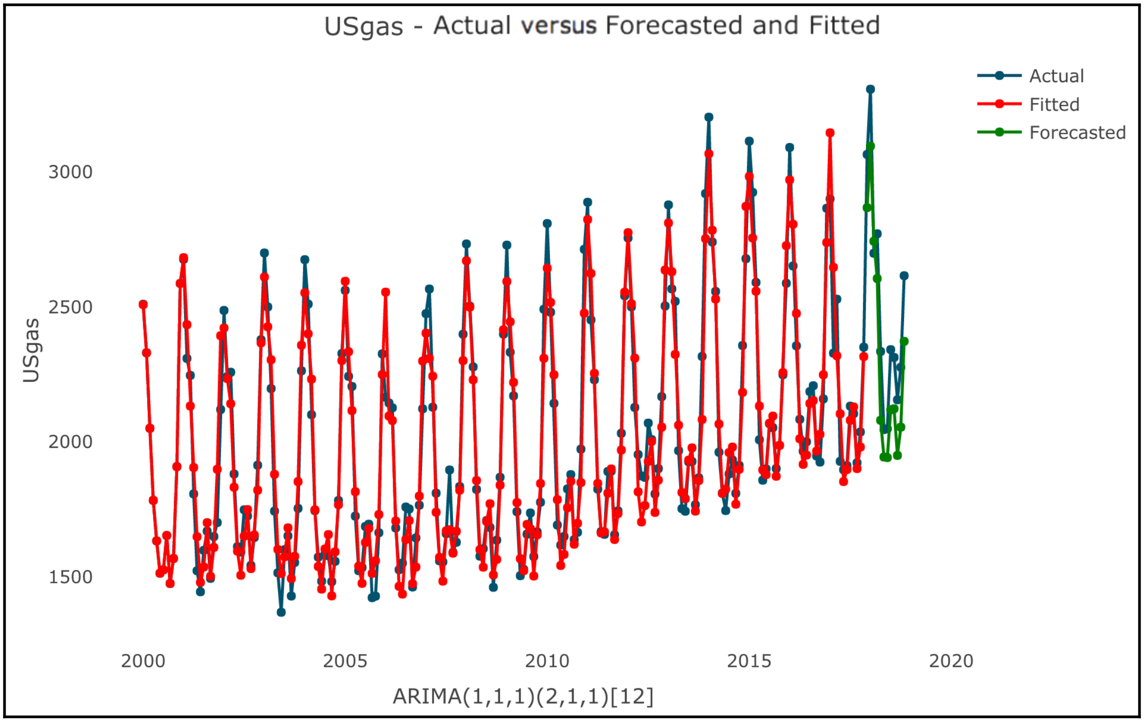

US Monthly Natural Gas consumption
Billion Cubic Feet
3000
2500
2000
1500
2000
2005
2010
2015
Year

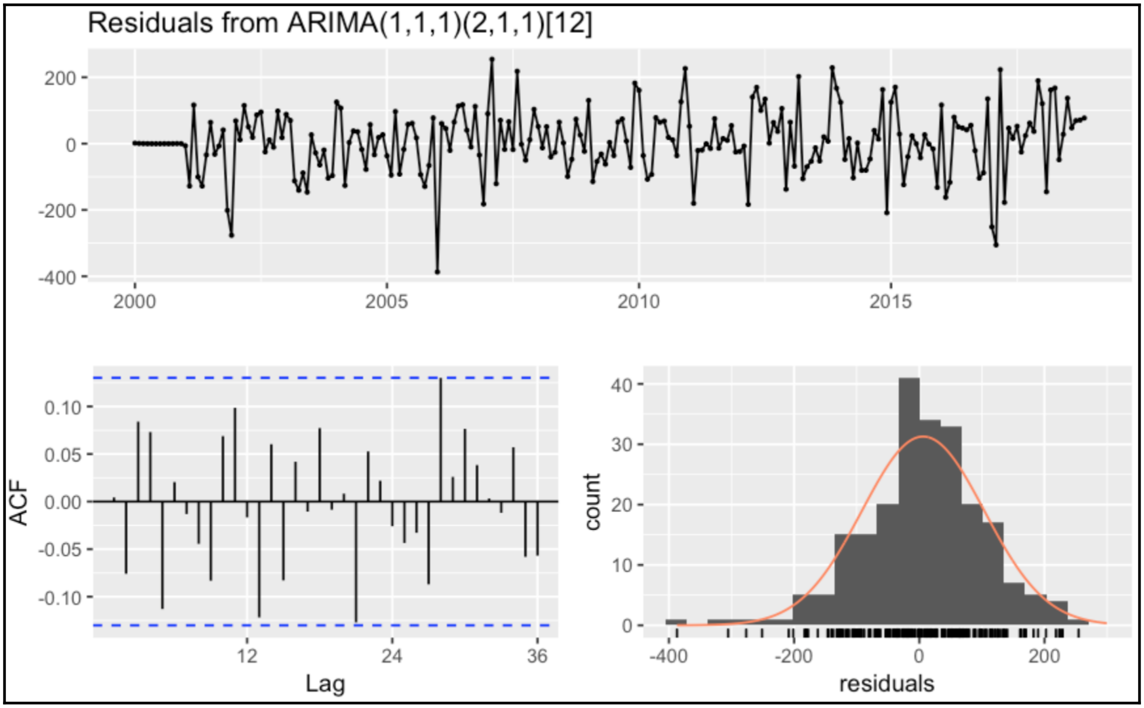

Series train
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0
1
2
3
4
5
Lag
Series train
Partial ACF
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
1
2
3
4
5
Lag

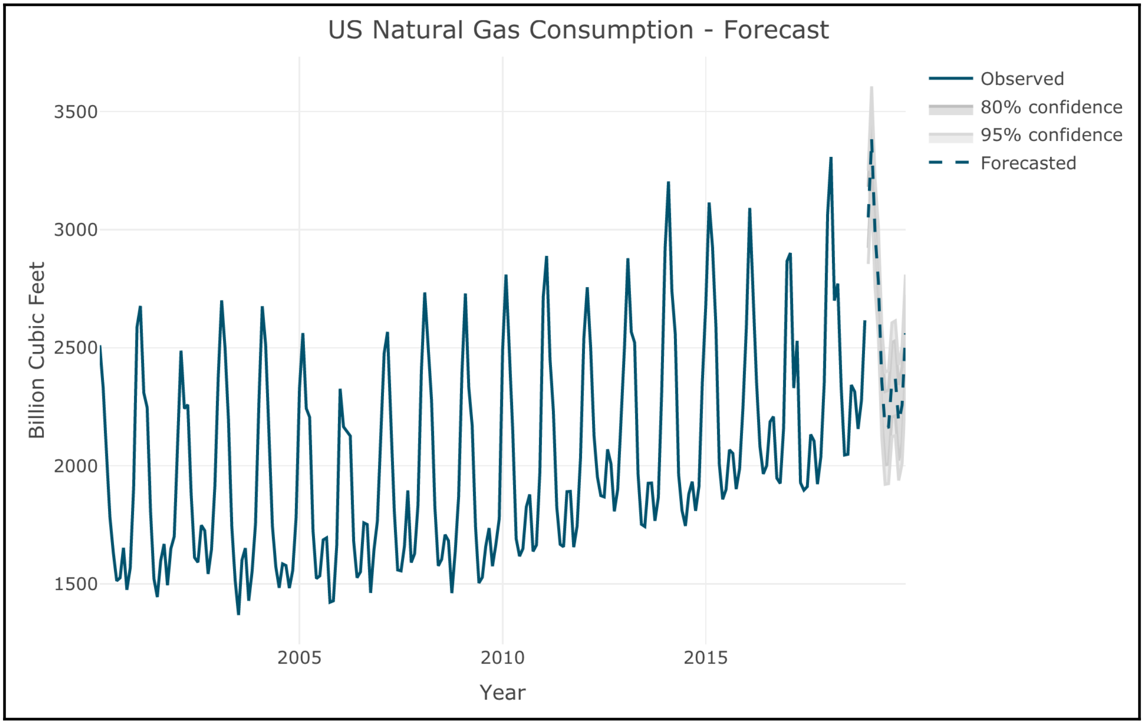

US Monthly Natural Gas consumption - First Seasonal Difference
400
300
200
100
0
−100
−200
−300
−400
−500
Billion Cubic Feet (First Difference)
2002
2004
2006
2008
2010
2012
2014
2016
Year

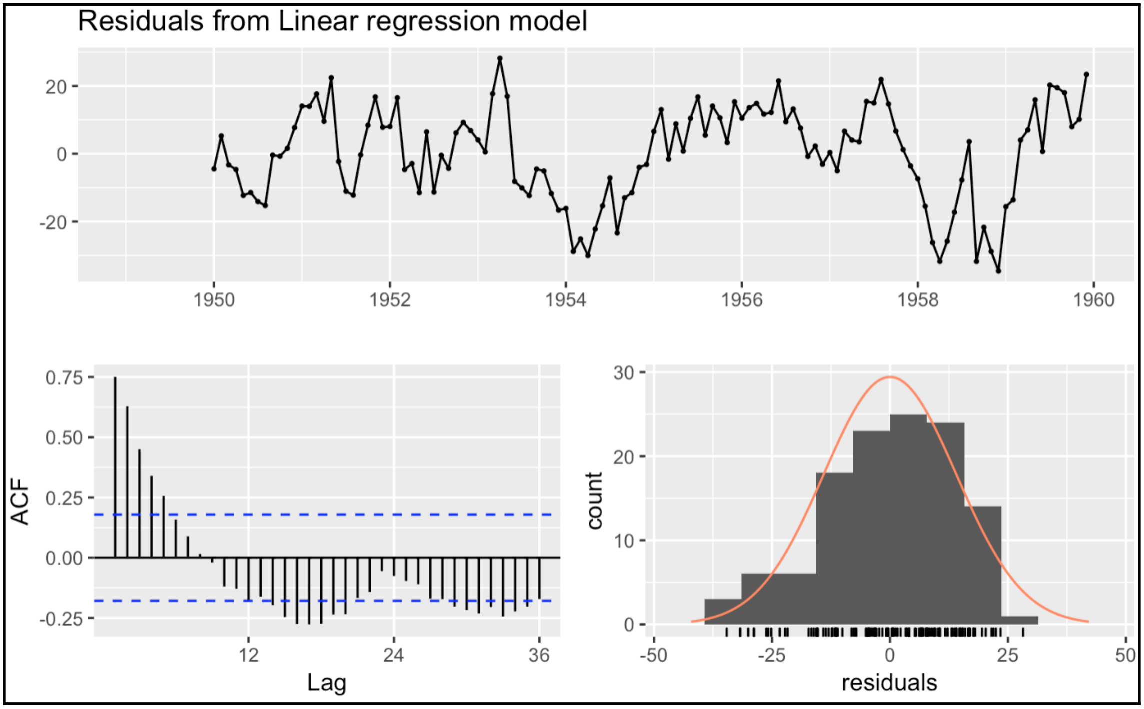

US Monthly Natural Gas consumption - First Seasonal and Non-Seasonal Differencing
Billion Cubic Feet (Difference)
800
600
400
200
0
−200
−400
−600
−800
2002
2004
2006
2008
2010
2012
2014
2016
Year

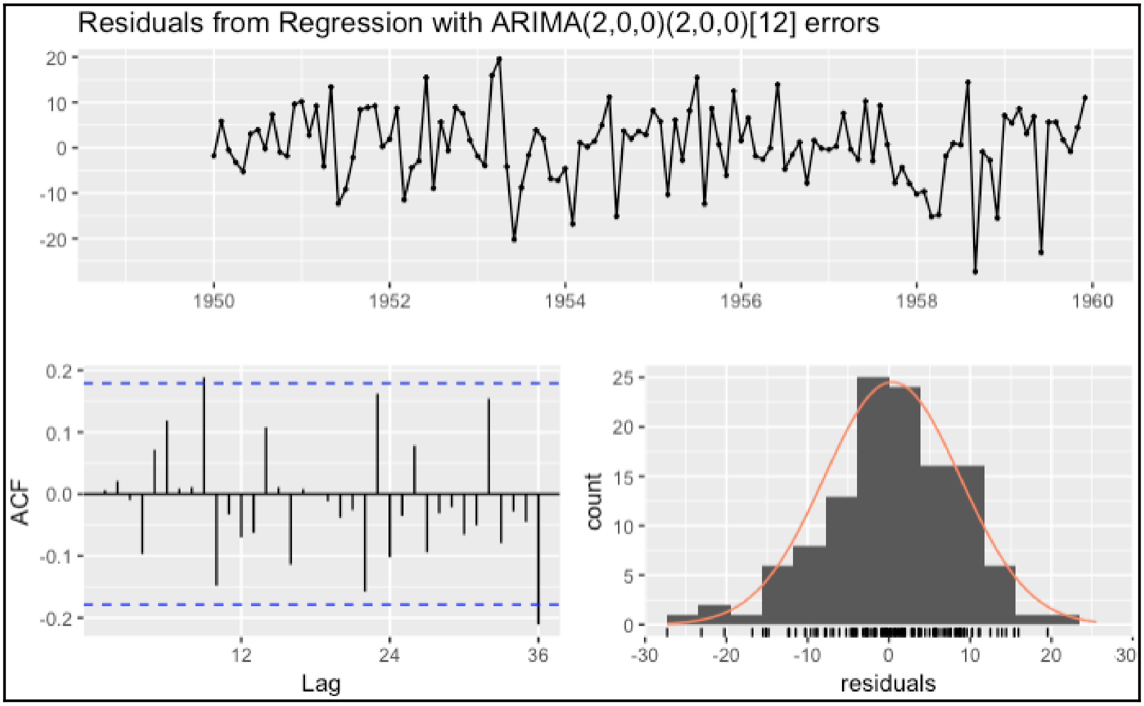

Series USgas_d12_1
ACF
Lag
Series USgas_d12_1
Partial ACF
Lag


USgas - Actual versus Forecasted and Fitted
Actual
Fitted
Forecasted
3000
2500
2000
1500
USgas
2000
2005
2010
2015
2020
ARIMA(1,1,1)(2,1,1)[12]


Residuals from ARIMA(1,1,1)(2,1,1)[12]
200
0
-200
-400
2000
2005
2010
2015
0.10
0.05
0.00
-0.05
-0.10
ACF
12
24
36
Lag
40
30
20
10
0
count
-400
-200
0
200
residuals

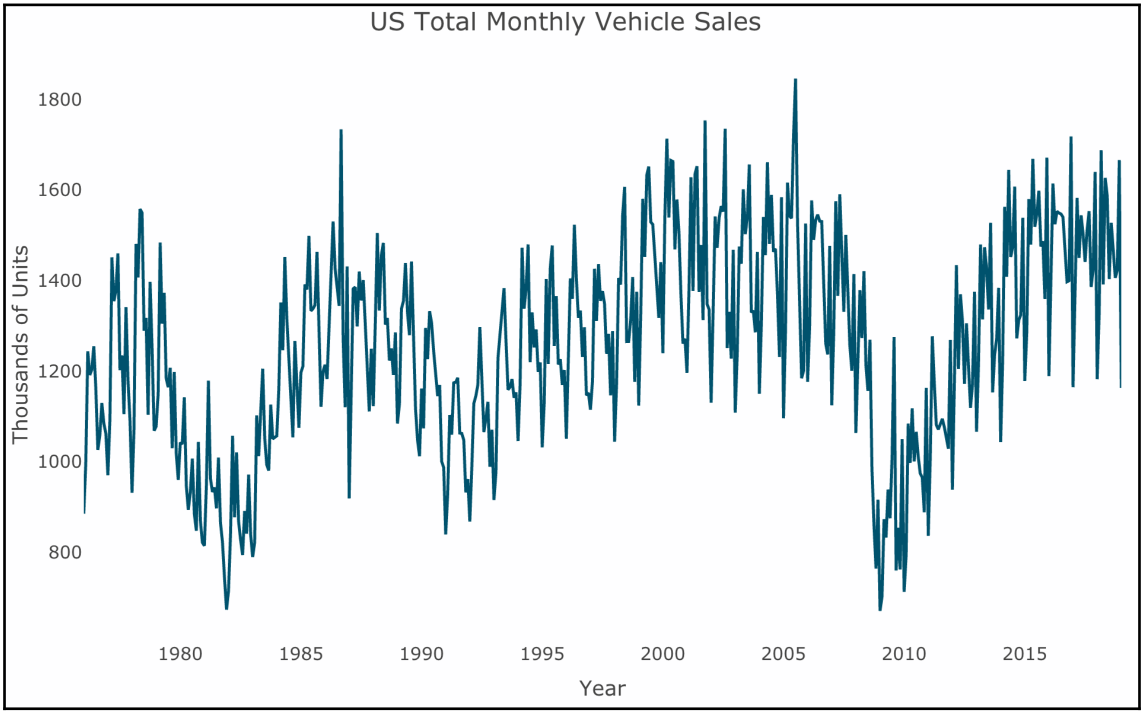

US Natural Gas Consumption - Forecast
Billion Cubic Feet
3500
3000
2500
2000
1500
2005
2010
2015
Year
Observed
80% confidence
95% confidence
Forecasted

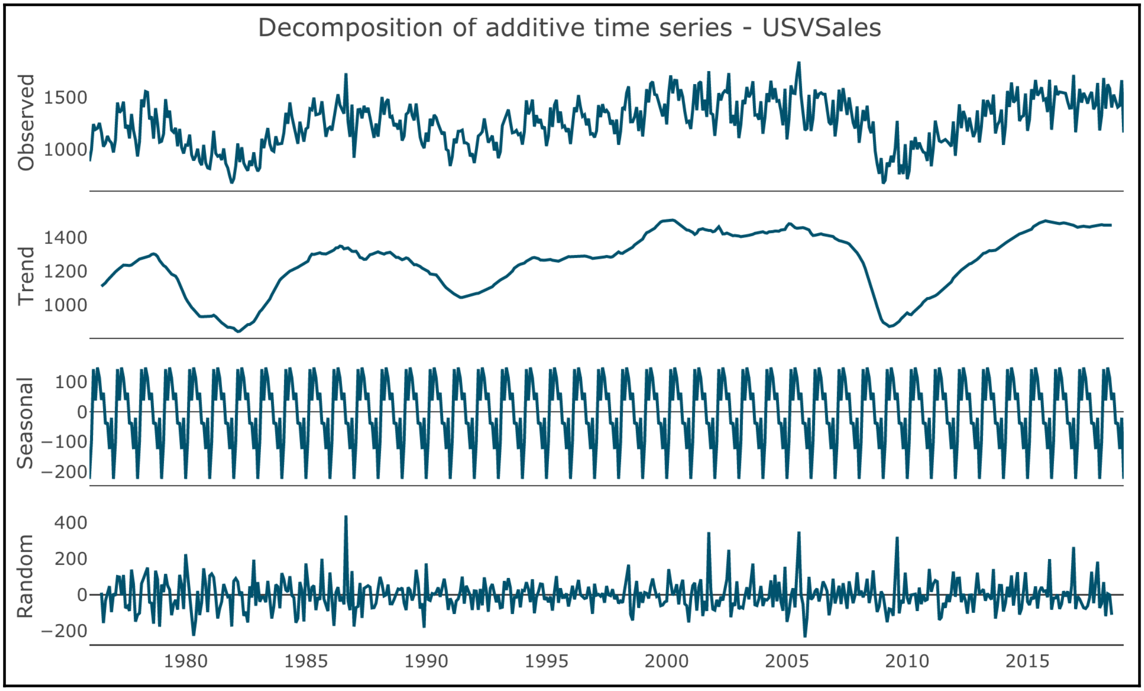

Residuals from Linear regression model
20
0
-20
1950
1952
1954
1956
1958
1960
ACF
0.75
0.50
0.25
0.00
-0.25
12
24
36
Lag
count
30
20
10
0
-50
-25
0
25
50
residuals

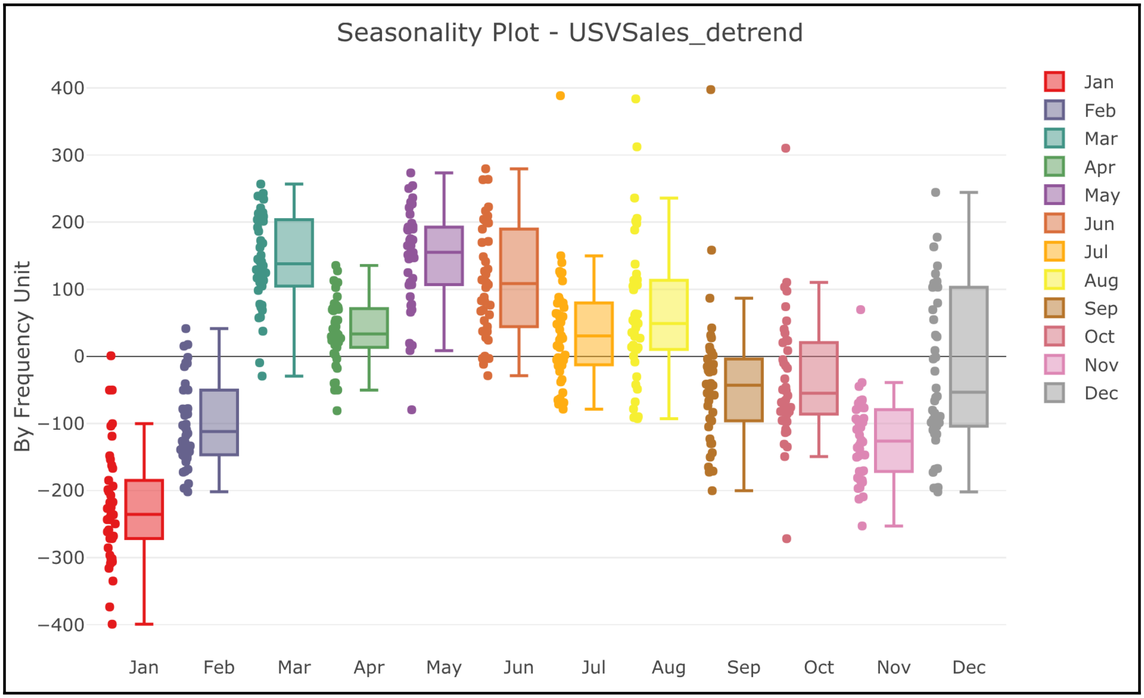


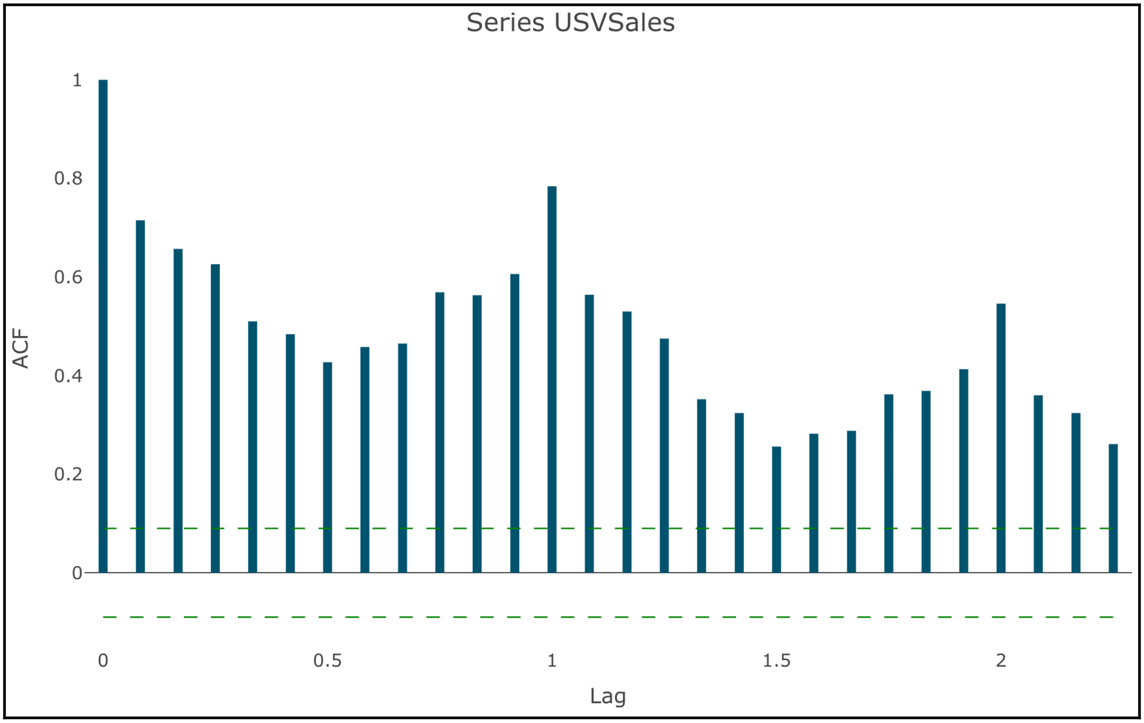
Residuals from Regression with ARIMA(2,0,0)(2,0,0)[12] errors
20
10
0
-10
-20
1950
1952
1954
1956
1958
1960
0.2
0.1
0.0
-0.1
-0.2
ACF
12
24
36
Lag
25
20
15
10
5
0
count
-30
-20
-10
0
10
20
30
residuals

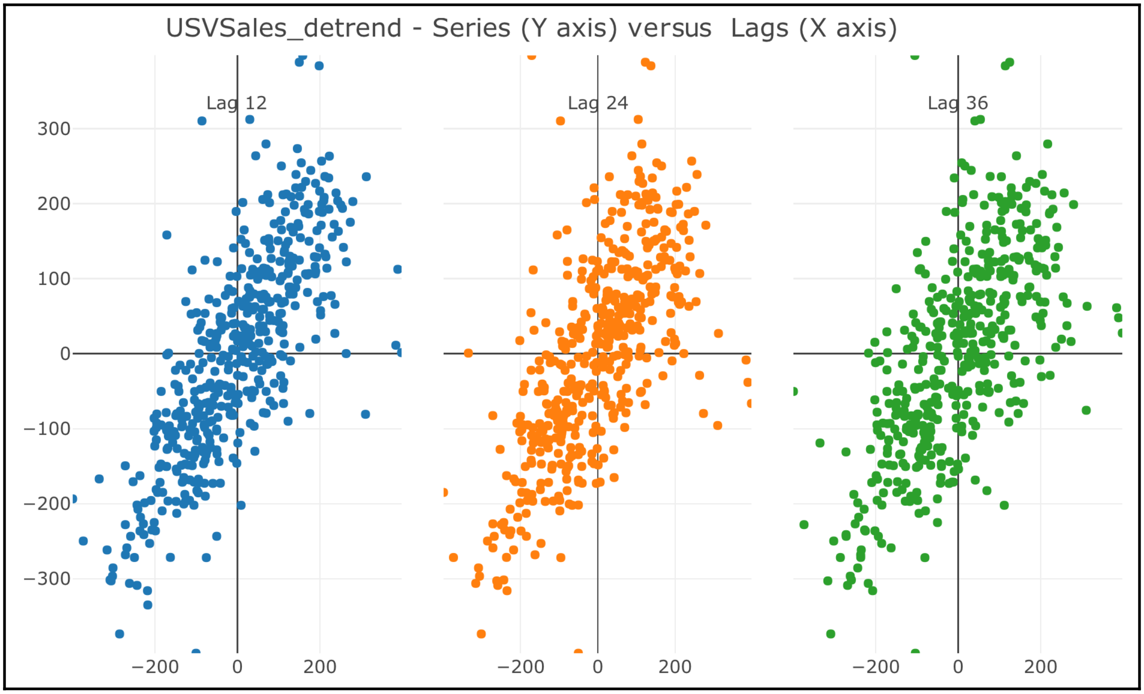
Chapter 12: Forecasting with Machine Learning Models

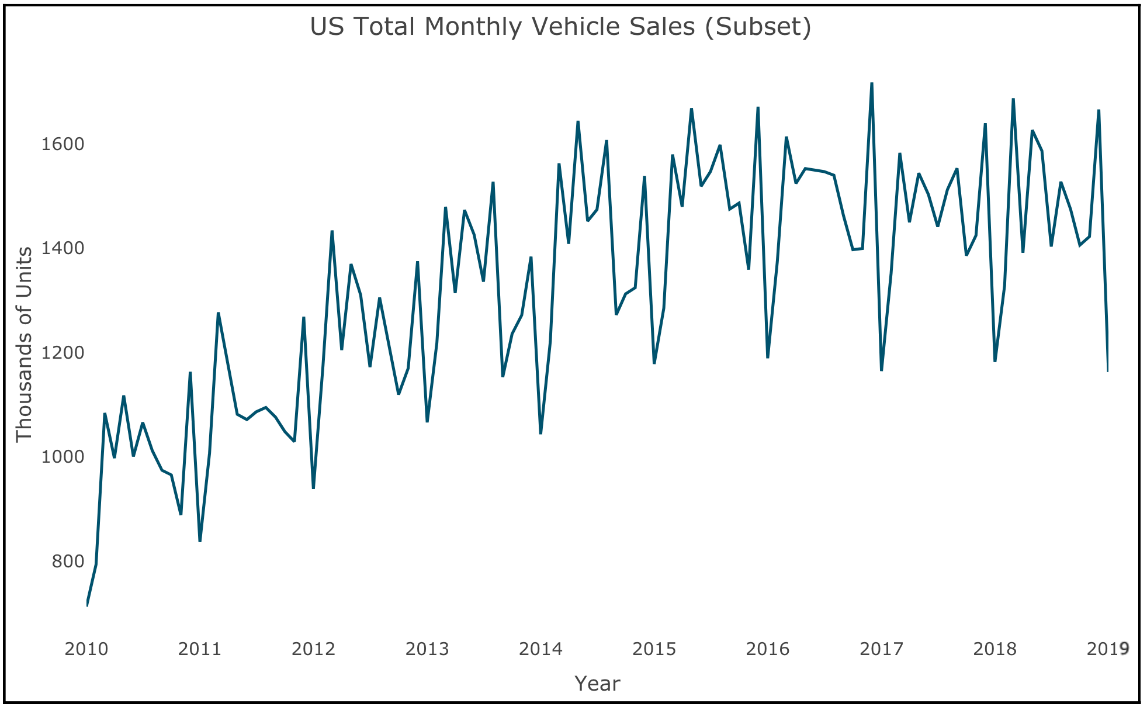


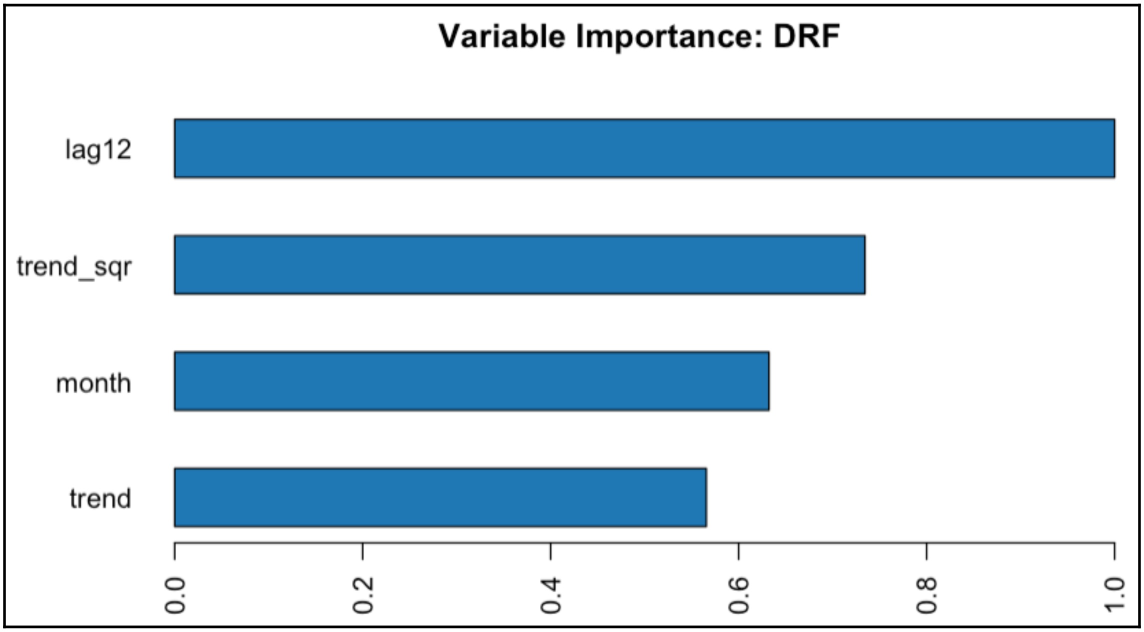

Decomposition of additive time series - USVSales
Observed
1500
1000
Trend
1400
1200
1000
Seasonal
100
0
−100
−200
Random
400
200
0
−200
1980
1985
1990
1995
2000
2005
2010
2015

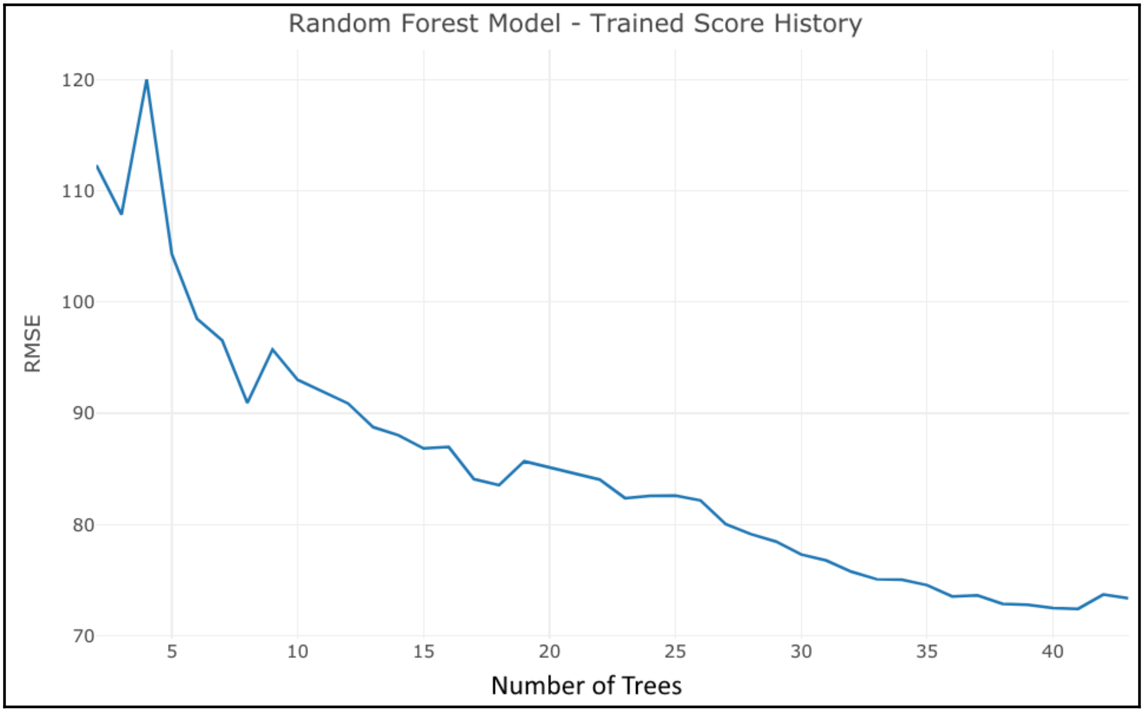

Seasonality Plot - USVSales_detrend
By Frequency Unit
400
300
200
100
0
−100
−200
−300
−400
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

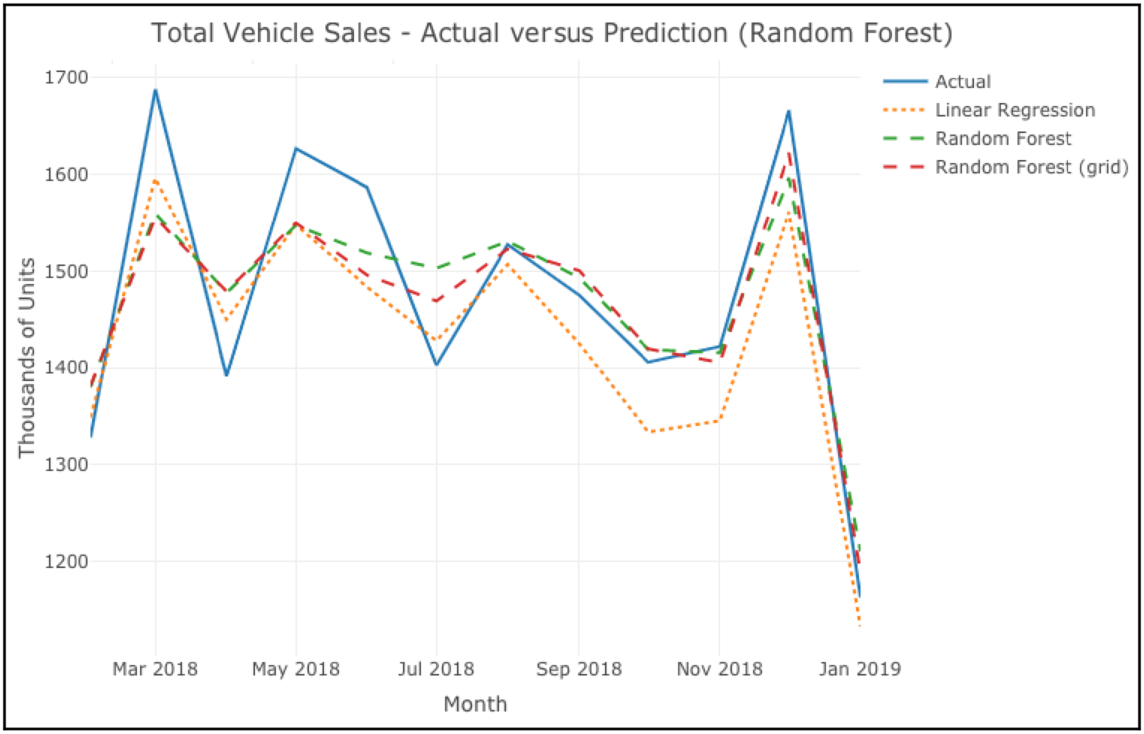

Series USVSales
ACF
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
Lag

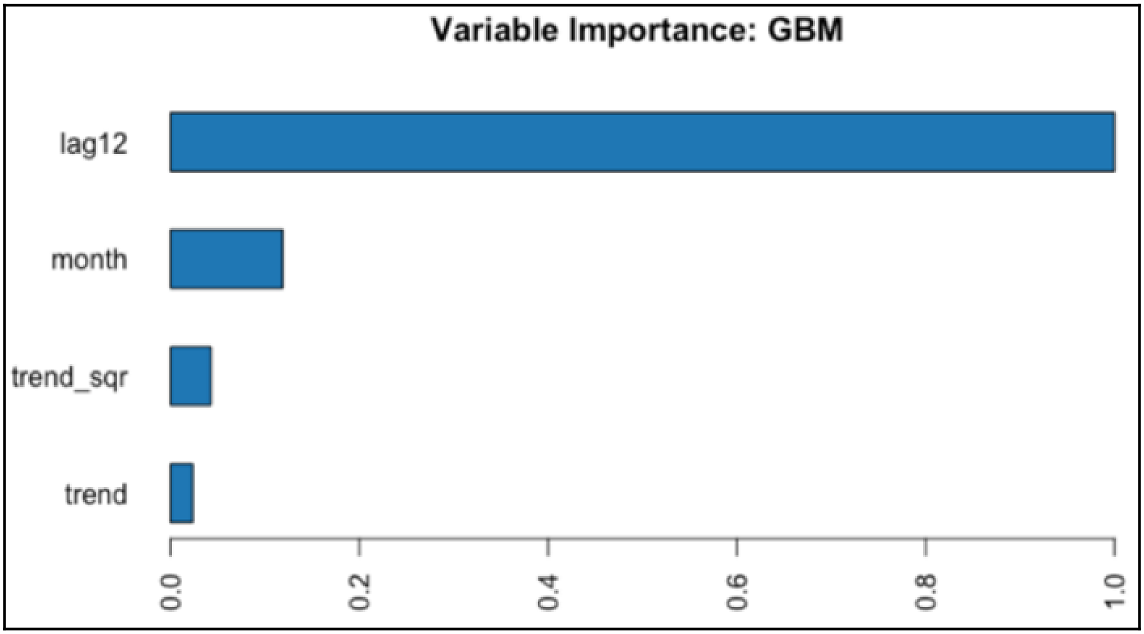

USVSales_detrend - Series (Y axis) versus Lags (X axis)
Lag 12
Lag 24
Lag 36
300
200
100
0
−100
−200
−300
−200
0
200

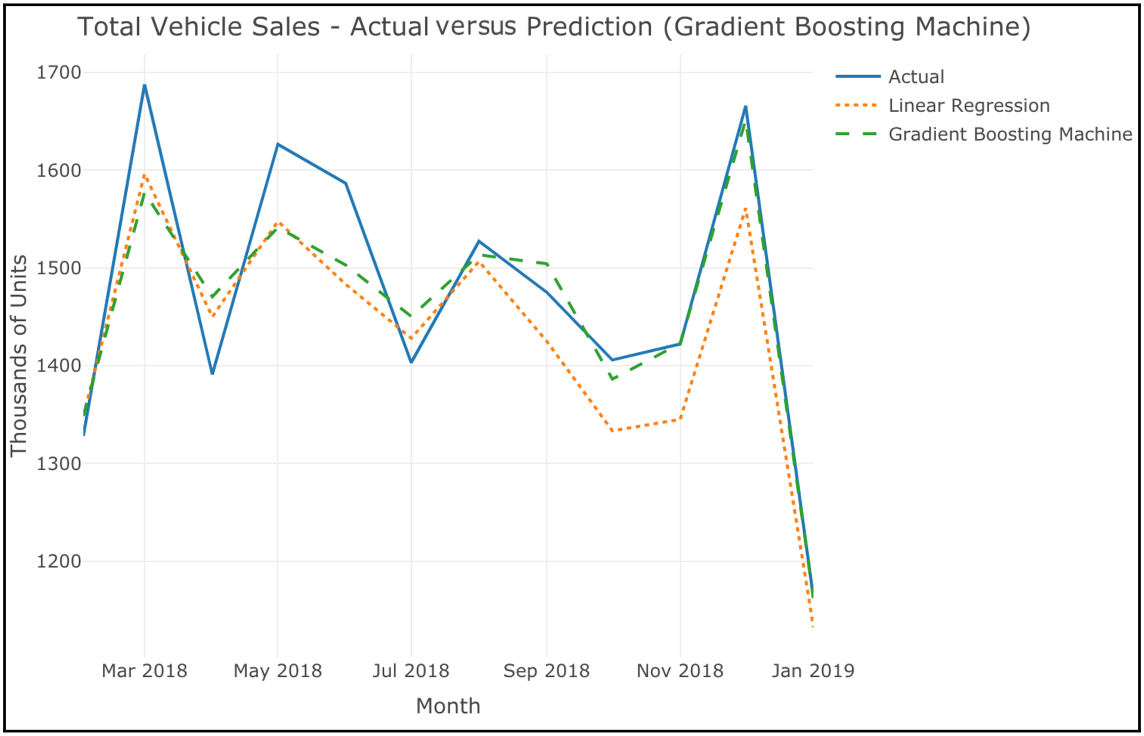

US Total Monthly Vehicle Sales (Subset)
Thousands of Units
1600
1400
1200
1000
800
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Year

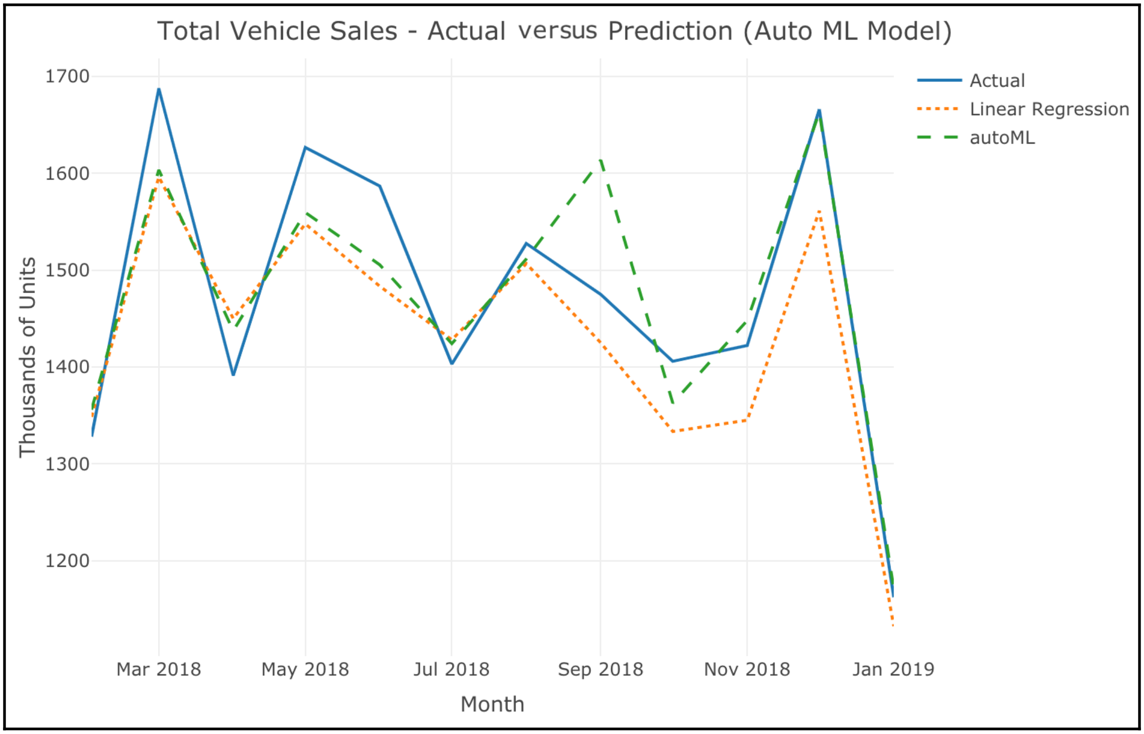

Variable Importance: DRF
lag12
trend_sqr
month
trend
0.0
0.2
0.4
0.6
0.8
1.0

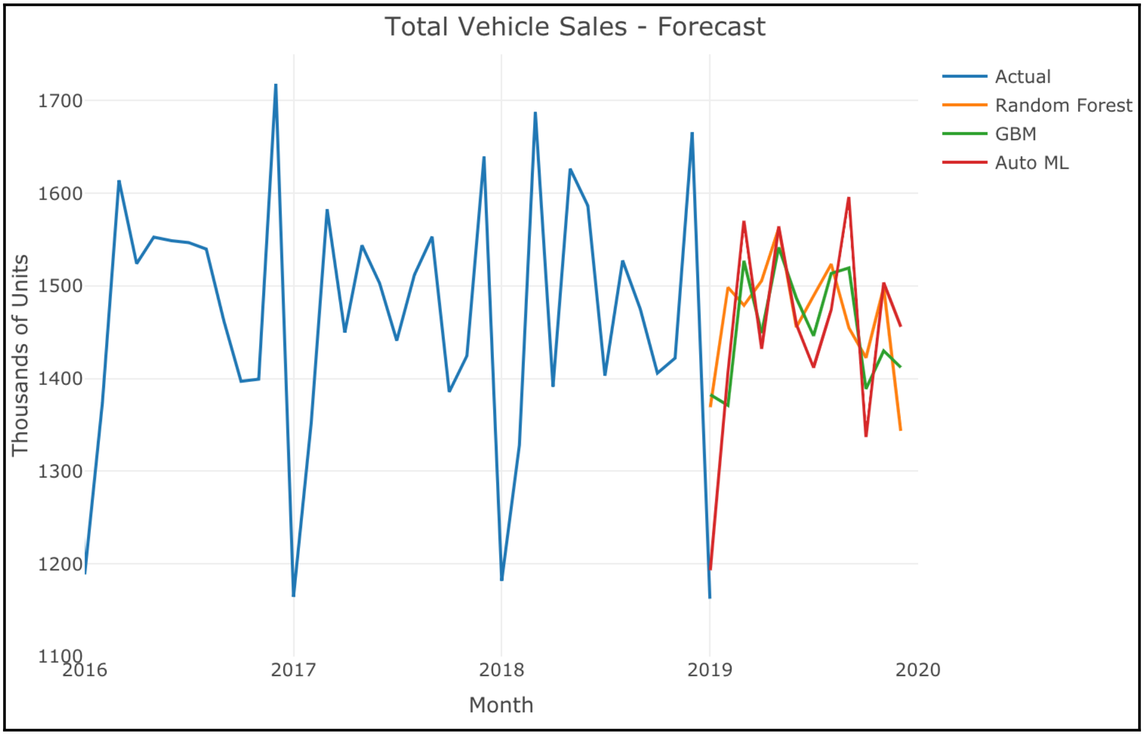

Random Forest Model - Trained Score History
RMSE
120
110
100
90
80
70
5
10
15
20
25
30
35
40
Number of Trees


Total Vehicle Sales - Actual versus Prediction (Random Forest)
Actual
Linear Regression
Random Forest
Random Forest (grid)
Thousands of Units
1700
1600
1500
1400
1300
1200
Mar 2018
May 2018
Jul 2018
Sep 2018
Nov 2018
Jan 2019
Month


Variable Importance: GBM
lag12
month
trend_sqr
trend
0.0
0.2
0.4
0.6
0.8
1.0


Total Vehicle Sales - Actual versus Prediction (Gradient Boosting Machine)
Actual
Linear Regression
Gradient Boosting Machine
Thousands of Units
1700
1600
1500
1400
1300
1200
Mar 2018
May 2018
Jul 2018
Sep 2018
Nov 2018
Jan 2019
Month


Total Vehicle Sales - Actual versus Prediction (Auto ML Model)
Thousands of Units
1700
1600
1500
1400
1300
1200
Mar 2018
May 2018
Jul 2018
Sep 2018
Nov 2018
Jan 2019
Month
Actual
Linear Regression
autoML


Total Vehicle Sales - Forecast
Actual
Random Forest
GBM
Auto ML
Thousands of Units
1700
1600
1500
1400
1300
1200
1100
2016
2017
2018
2019
2020
Month